Accelerating functional gene discovery in osteoarthritis

Natalie C. Butterfield[1], Katherine F. Curry[1], Julia Steinberg[2,3,4], Hannah Dewhurst[1], Davide Komla-Ebri[1], Naila S. Mannan[1], Anne-Tounsia Adoum[1], Victoria D. Leitch[1], John G. Logan[1], Julian A. Waung[1], Elena Ghirardello[1], Lorraine Southam[2,3], Scott E. Youlten[5], J, Mark Wilkinson[6,7], Elizabeth A. McAninch[8], Valerie E. Vancollie[3], Fiona Kussy[3], Jacqueline K. White[3,9], Christopher J. Lelliott[3], David J. Adams[3], Richard Jacques[10], Antonio C. Bianco[11], Alan Boyde[12], Eleftheria Zeggini[2,3], Peter I. Croucher[5], Graham R. Williams[1,13*] and J. H. Duncan Bassett[1,13*]

[1]Molecular Endocrinology Laboratory, Department of Metabolism, Digestion and Reproduction, Imperial College London, London W12 0NN, UK
[2]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany
[3]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
[4]Cancer Council NSW, Sydney, New South Wales 2000, Australia
[5]The Garvan Institute of Medical Research and St. Vincent's Clinical School, University of New South Wales Medicine, Sydney, New South Wales 2010, Australia
[6]Department of Oncology and Metabolism, University of Sheffield, Sheffield S10 2RX, UK
[7]Centre for Integrated Research into Musculoskeletal Ageing and Sheffield Healthy Lifespan Institute, University of Sheffield, Sheffield S10 2TN, UK
[8]Division of Endocrinology and Metabolism, Rush University Medical Center, Chicago, IL 60612, USA
[9]The Jackson Laboratory, Bar Harbor, ME 04609, USA
[10]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield S1 4DA, UK
[11]Section of Adult and Pediatric Endocrinology, Diabetes & Metabolism, Department of Medicine, University of Chicago, Chicago, IL 60637, USA
[12]Dental Physical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

[13]These authors contributed equally
[*]Corresponding authors: Graham R. Williams (graham.williams@imperial.ac.uk) and J. H. Duncan Bassett (d.bassett@imperial.ac.uk)

Abstract

Osteoarthritis causes debilitating pain and disability, resulting in a considerable socioeconomic burden, yet no drugs are available that prevent disease onset or progression. Here, we develop, validate and use rapid-throughput imaging techniques to identify abnormal joint phenotypes in randomly selected mutant mice generated by the International Knockout Mouse Consortium. We identify 14 genes with functional involvement in osteoarthritis pathogenesis, including the homeobox gene *Pitx1*, and functionally characterize 6 candidate human osteoarthritis genes in mouse models. We demonstrate sensitivity of the methods by identifying age-related degenerative joint damage in wild-type mice. Finally, we phenotype previously generated mutant mice with an osteoarthritis-associated polymorphism in the *Dio2* gene by *CRISPR/Cas9* genome editing and demonstrate a protective role in disease onset with public health implications. This expanding resource of mutant mice will accelerate functional gene discovery in osteoarthritis and offer drug discovery opportunities for this common, incapacitating chronic disease.

Introduction

Osteoarthritis is the commonest cause of joint destruction, pain and disability. Joint replacement for end-stage disease remains the only treatment, leading to an escalating healthcare crisis in our obese and ageing society. Osteoarthritis is a complex trait and the 86 reported genome-wide associated loci explain only a small proportion of its heritability, which is estimated between 40-70%[1, 2, 3].

Osteoarthritis is characterized by articular cartilage damage and loss, together with structural abnormalities of subchondral bone and low-grade chronic joint inflammation. It is unknown which of these processes trigger disease or which represent secondary responses to joint destruction[4]. It is also uncertain whether the pathogenesis of osteoarthritis reflects an abnormal response to injury involving defective stem cell recruitment and abnormal cell proliferation, differentiation, metabolism, apoptosis and senescence[5, 6].

Chondrocytes in healthy articular cartilage are resistant to terminal differentiation whereas they revert to a developmental program following injury, in which they proliferate and undergo hypertrophic differentiation with accelerated cartilage mineralization[7]. Osteoarthritis pathogenesis involves cross-talk between the synovium, articular cartilage and subchondral bone[8], although the timing of bone remodeling relative to cartilage degradation remains uncertain[9]. Nevertheless, increases in apoptotic and senescent chondrocytes are triggered by processes including endoplasmic reticulum stress[10]. Senescent cells express a secretory phenotype that contributes to inflammation, vascular invasion[11] and cartilage breakdown via key pathways that stimulate matrix metalloproteases[12] and aggrecan-specific proteinases[13].

Despite the profound clinical and socio-economic impacts of osteoarthritis, our understanding of its genetic basis is in its infancy. We hypothesize that accelerating gene discovery in osteoarthritis will increase understanding of joint physiology and disease pathogenesis and facilitate identification of drug targets that prevent or delay joint destruction. Studies of extreme

phenotypes in humans have underpinned identification of the molecular basis of single gene disorders and mechanisms of complex disease and resulted in new treatments[14, 15]. Analogous to our gene discovery studies in osteoporosis[16, 17, 18, 19, 20], we propose that a joint-specific extreme phenotype screen in mutant mice will accelerate functional gene discovery in osteoarthritis.

Mutant mice are generated at the Sanger Institute as part of the International Mouse Phenotyping Consortium (IMPC). Mice undergo broad phenotyping using the International Mouse Phenotyping Resource of Standardized Screen (IMPReSS) that is completed at 16 weeks of age when tissues are harvested for further analysis. In the Origins of Bone and Cartilage Disease (OBCD) Project we collaborate with IMPC and receive knee joints for analysis. Rapid-throughput phenotyping of the mouse knee requires quantitative imaging; this presents a complex and unsolved challenge that relates to anatomical size, three-dimensional complexity, image resolution and the necessity to maintain joints in their native fully-hydrated state.

Here we present the invention, optimization, validation and application of a rapid-throughput multimodality imaging pipeline to phenotype the mouse knee. We analyze 50 randomly selected mouse lines, identifying seven (14%) with markedly abnormal phenotypes. A systematic prioritization strategy identifies seven further lines, resulting in 14 genes (28%) with evidence for a functional role in osteoarthritis pathogenesis. The four leading candidates are *Pitx1*, *Bhlhe40*, *Sh3bp4* and *Unk*. We interrogate the database of joint phenotypes from randomly selected mouse lines with 409 genes differentially expressed in human osteoarthritis cartilage. This results in an enriched yield of abnormal joint phenotypes in six (75%) of eight lines for which data are available, including *Unk*. We then apply the pipeline to characterize the early features of age-related joint degeneration in one-year old mice and demonstrate its sensitivity to detect disease onset as well as surgically provoked late-stage disease, paving the way for application to analysis of drug intervention studies. Finally, we phenotype

previously generated *CRISPR/Cas9* mutant mice with a Thr92Ala polymorphism in the *Dio2* gene that is orthologous to the human variant associated with osteoarthritis susceptibility. The Ala92 allele confers protection against early-onset osteoarthritis, challenging current understanding with implications for public health.

Results

Invention and optimization of imaging methods

We established a rapid-throughput joint phenotyping pipeline (OBCD joint pipeline), which applies three complementary imaging approaches to characterize features of osteoarthritis in mutant mice that include articular cartilage damage and loss, together with abnormalities of subchondral bone structure and mineralization (Figure 1). Iodine contrast-enhanced micro-computerized tomography (ICEμCT) was developed to determine articular cartilage volume (Cg.V), median thickness (Median Cg.Th), maximum thickness (Max Cg.Th), subchondral bone volume/tissue volume (SC BV/TV), trabecular thickness (SC Tb.Th), trabecular number (SC Tb.N) and tissue mineral density (SC TMD). Joint surface replication (JSR) was invented and optimized to quantify articular cartilage surface damage. Subchondral bone X-ray microradiography (scXRM) was developed from previous protocols[21, 22] to determine subchondral bone mineral content (SC BMC) (Figures 1, 2 and Supplementary Figure 1).

Validation of imaging methods

The OBCD joint pipeline methods were validated by comparison with Osteoarthritis Research Society International (OARSI) histological scoring[23] of knees from 22-week-old wild-type (WT) mice, 12 weeks after destabilization of the medial meniscus (DMM)[13, 24]. One cohort of WT mice (n=16) was phenotyped using ICEμCT, JSR and scXRM in the OBCD joint pipeline, and a second cohort (n=11) was analyzed by OARSI scoring (Figure 1). DMM surgery resulted in decreased Cg.V and Max Cg.Th with a marked increase in Cg. damage (Figure 3). This cartilage destruction was accompanied by increased SC BV/TV, increased SC Tb.Th, decreased SC Tb.N and increased SC BMC (Figure 3 and Supplementary Figure 2). These abnormalities were consistent with extensive cartilage damage, synovitis and osteophyte formation demonstrated by OARSI-scored histology together with ICEμCT and JSR (Figure 3 and Supplementary Figure 2, Supplementary Figure 3 and Supplementary Data 1, Supplementary Data 2).

OARSI analysis was also undertaken on three joints that were phenotyped in the OBCD joint pipeline and found to have features indicating mild, intermediate and severe osteoarthritis following DMM surgery. OARSI scoring was concordant with the severity of abnormalities identified by ICEμCT, JSR and scXRM, thus validating the use of these imaging modalities to characterize osteoarthritis (Supplementary Figure 4).

Definition of WT reference ranges

Phenotype datasets for all parameters were obtained from 100 16-week-old male WT mice. Reference ranges, coefficients of variation, estimates of skewness and kurtosis, normality, repeatability and power calculations were determined (Supplementary Data 3). The reference range for each parameter was defined as either (i) the mean +/- 2.0 standard deviations for normally-distributed data, or (ii) the median and 2.5^{th}-97.5^{th} percentile range for non-normally distributed data (Supplementary Figure 1).

Identification of osteoarthritis genes

To identify genes that cause osteoarthritis, we used the OBCD joint pipeline to analyze knees from 16-week-old male mice (n=3-7) from 50 randomly selected mutant lines generated in an identical *C57BL6/N;C57BL6/NTac* genetic background (Supplementary Data 4).

Rigorous statistical approaches were used to determine which lines displayed robust outlier phenotypes (Figure 1, Supplementary Data 4 and 5). Lines with outlier phenotypes were identified: (i) when the mean value for an individual parameter was outside the WT reference range; (ii) by using the Wilcoxon rank sum statistical test to analyze all individual phenotype measurements rather than only mean values, with a Bonferroni multiple-testing correction for the effective number of tests (Supplementary Data 5); and (iii) by calculation of Mahalanobis distances to ensure that lines with significantly abnormal phenotypes were not overlooked

when they resulted from simultaneous but smaller variances in multiple phenotype parameters. 17 lines (34%) had at least one abnormal phenotype parameter outside the reference range, 12 (24%) of which had increased cartilage surface damage. 18 lines (36%) were outliers after statistical analysis using the Wilcoxon test with Bonferroni correction, and 10 (20%) were outliers following Mahalanobis analysis. Overall, 25 individual lines (50%) had an abnormal joint phenotype based on these criteria (Figure 1). Increased body weight was not present in any mutant line (Supplementary Data 4). Furthermore, body weight did not correlate with any joint phenotype parameter in wild-type mice (P<0.05, Spearman correlation, Supplementary Data 6).

The effect of gene deletion on phenotype severity was assessed by scoring whether joint abnormalities were (i) reference range outliers, (ii) outliers based on the Bonferroni-corrected Wilcoxon test, (iii) outliers after Mahalanobis analysis, and by considering whether joint pathology included abnormalities of cartilage morphology, cartilage integrity and/or subchondral bone structure (Supplementary Data 7). Of the 25 lines with an outlier phenotype, the top seven ranked lines with the most severe abnormalities were a cluster of 6 microRNAs (*ClusterXN1: miR-106a, miR-18b, miR-19b-2, miR-20b, miR-92-2, miR-363*), paired like homeodomain 1 (*Pitx1*), fibroblast growth factor binding protein 1 (*Fgfbp1*), Josephin domain containing 1 (*Josd1*), SH3 domain binding protein 4 (*Sh3bp4*), unkempt family zinc finger (*Unk*), and zinc finger protein 341 (*Zfp341*) (Figure 1, Supplementary Data 7).

Prioritization of 25 mouse lines with outlier phenotypes

An informatics strategy was used to investigate biological plausibility and prioritize the 25 candidate genes based on (i) additional skeletal consequences of gene deletion, (ii) gene expression in skeletal cells and tissues, (iii) association with human disease, and (iv) structured literature searching. Each criterion was assigned a score and scores were summed to rank genes in order of priority.

Additional skeletal consequences of gene deletion were investigated by: (i) identifying whether mutant mice had abnormalities of bone structure and strength[16], (ii) IMPC IMPReSS phenotype screening, (iii) analysis of the skeleton in embryos from lines in which homozygous gene deletion was lethal or resulted in sub-viability[25], and (iv) determining whether mutations of the same gene had skeletal abnormalities identified in Mouse Genome Informatics (MGI) databases or the published literature. The top ranked genes with major effects on skeletal phenotype were *Pitx1*, basic helix-loop-helix family member e40 (*Bhlhe40*), SR-related CTD associated factor 11 (*Scaf11*) and SMG9 nonsense mediated mRNA decay factor (*Smg9*) (Figure 1, Supplementary Data 7).

Gene expression in skeletal cells and tissues was investigated by interrogation of (i) MGI and BioGPS[26] databases, and (ii) transcriptome datasets from human chondrocytes and cartilage[27, 28], mouse and human osteoblasts[26, 28], mouse osteocytes[19], and mouse and human osteoclasts[26, 28]. The top ranked genes with expression enriched in cartilage and bone compared to non-skeletal tissues were *Bhlhe40*, *Pitx1*, collagen type IV alpha 2 chain (*Col4a2*) and *Sh3bp4* (Figure 1, Supplementary Data 7).

Association with human disease was investigated by searching MGI and Online Mendelian Inheritance in Man (OMIM) databases to determine whether mutations in any of the 25 genes resulted in monogenic diseases affecting the skeleton. To determine whether any gene loci were associated with arthritis or other skeletal phenotypes, we interrogated the European Bioinformatics Institute GWAS catalogue. The top ranked genes associated with arthritis and skeletal disease in humans were *Pitx1*, coiled-coil domain containing 6 (*Ccdc6*), HECT and RLD domain containing E3 ubiquitin protein ligase family member 1 (*Herc1*), nebulette (*Nebl*), *Sh3bp4* and *Zfp341* (Figure 1, Supplementary Data 7).

Structured literature searching of PubMed and Google Scholar databases identified *Bhlhe40*, *Col4a2* and *Pitx1* as the top ranked candidate genes with publications related to arthritis or skeletal cell biology (Figure 1, Supplementary Data 7).

Together, consideration of additional skeletal phenotypes, gene expression, association with human disease and the published literature prioritized 10 lines (*Pitx1*, *Bhlhe40*, *Scaf11*, *Smg9*, *Col4a2*, *Sh3bp4*, *Ccdc6*, *Herc1, Nebl* and *Zfp341*). 11 (44%) of the 25 randomly selected candidates with outlier phenotypes (*Arhgap30*, *Arrdc5*, *ClusterXN1*, *Cpgi81*, *Gsdme*, *Hpf1*, *Josd1*, *Mkrn2*, *Scaf11*, *Slamf9*, *Smg9*) had no prior links to osteoarthritis or skeletal biology based on association with human disease and structured literature searching, whereas only two (20%) of the 10 prioritized candidates (*Scaf11*, *Smg9*) had no prior association (Figure 1, Supplementary Data 7).

Combining findings from informatics prioritization with the seven lines with the most severe joint phenotypes (*ClusterXN1*, *Pitx1*, *Fgfbp1*, *Josd1*, *Sh3bp4*, *Unk*, *Zfp341*) identified 14 genes with strong evidence from independent sources for a functional role in osteoarthritis pathogenesis (*Pitx1*, *Bhlhe40*, *Scaf11*, *Smg9*, *Col4a2*, *Sh3bp4*, *Ccdc6*, *Herc1, Nebl*, *Zfp341, ClusterXN1*, *Fgfbp1*, *Josd1*, and *Unk*). The four genes with the strongest overall evidence were *Pitx1*, *Bhlhe40*, *Sh3bp4* and *Unk* (Figure 1, Supplementary Data 7).

Severe early onset osteoarthritis in Pitx1$^{+/-}$ mice

Pitx1 is a homeobox transcription factor required for patterning, and specification of hindlimb morphology[29, 30]. Homozygous *Pitx1$^{-/-}$* mutations cause post-natal lethality with gross morphological abnormalities affecting the hindlimb skeleton[29, 30]. Genomic rearrangements at the human *PITX1* locus result in homeotic arm-to-leg transformations in Liebenberg syndrome (OMIM 186550), and misexpression of *Pitx1* in the mouse forelimb recapitulates this phenotype[31]. Haploinsufficiency for *Pitx1* in mice and humans results in clubfoot and other leg malformations, demonstrating hindlimb development is sensitive to *Pitx1* gene dosage[32].

Heterozygous *Pitx1*$^{+/-}$ mice had the most severe phenotype observed in this study, with extensive joint damage affecting both compartments of the knee including decreased articular cartilage volume and thickness, and increased areas of complete loss or severe erosion of articular cartilage. The *Pitx1*$^{+/-}$ phenotype was as severe as the extensive joint damage observed in 22-week-old WT mice 12 weeks after DMM surgery (Figure 3 and Supplementary Figure 2a and c). Osteophytes were detected in 5/7 *Pitx1*$^{+/-}$ mice (Supplementary Data 8). Accordingly, histology confirmed severe osteoarthritis with extensive vertical clefts in articular cartilage and areas of erosion over 50-75% of the articular surface to the calcified cartilage beneath (Figure 3, Supplementary Figure 2c and Supplementary Data 4).

Early onset osteoarthritis in Bhlhe40*$^{-/-}$ *and Sh3pb4*$^{-/-}$ *mice

Bhlhe40$^{-/-}$ mice had decreased median and maximum Cg.Th affecting both compartments of the knee. Further evidence of spontaneous damage included vertical clefts in the articular cartilage surface extending down to the layer of chondrocytes immediately below the superficial layer, together with some loss of surface lamina (Figure 4, Supplementary Figure 5 and Supplementary Data 4). No osteophytes were detected in *Bhlhe40*$^{-/-}$ mice (Supplementary Data 8). *Bhlhe40* is a widely expressed transcription factor involved in regulation of cell proliferation, differentiation, apoptosis and senescence[33]. *Bhlhe40* expression is enriched in skeletal tissues (Supplementary Data 7), specifically in proliferating and differentiating chondrocytes during endochondral ossification[34], and is elevated in response to hypoxia[35], bone morphogenetic protein-2 (BMP2) and transforming growth factor-β (TGFβ) but suppressed by parathyroid hormone[36]. *Bhlhe40* stimulates terminal chondrocyte and osteoblast differentiation[36, 37], and has been implicated in bone loss in chronic periodontitis[38]. Deletion of *Bhlhe40* in mice results in increased bone mineral content and density (Supplementary Data 7), but a role in osteoarthritis pathogenesis has not been postulated and *BHLHE40* has not been associated with osteoarthritis in GWAS.

Overall, young adult *Bhlhe40*$^{-/-}$ mice display thinning of articular cartilage, suggesting a role in disease onset that involves the key hypoxia, BMP2, TGFβ and PTH signalling pathways[39, 40, 41, 42].

Sh3bp4$^{-/-}$ mice had decreased Cg.V and median and maximum Cg.Th, together with increased articular cartilage surface damage affecting the lateral compartment. Histology revealed moderate osteoarthritis with articular cartilage fibrillations and vertical clefts over <25% of the articular surface extending to the calcified cartilage beneath (Figure 4, Supplementary Figure 5 and Supplementary Data 4). No osteophytes were detected in *Sh3bp4*$^{-/-}$ mice (Supplementary Data 8). *Sh3bp4* is a poorly characterized SH3 domain binding protein involved in transferrin receptor (TfR) internalization[43], fibroblast growth factor receptor (FGFR) trafficking[44], mammalian target of rapamycin (mTOR) signalling[45] and inhibition of the Wnt pathway[46]. *SH3BP4* has not been associated with osteoarthritis in GWAS.

In summary, *Sh3bp4*$^{-/-}$ mice display loss of articular cartilage and moderate cartilage damage. *Sh3bp4* thus represents an osteoarthritis susceptibility gene, the deletion of which accelerates joint damage. Its role in disease pathogenesis may involve the key TfR, FGFR, mTOR and Wnt signalling pathways, which regulate bone and cartilage homeostasis and tissue repair[47, 48, 49, 50].

Additional applications of the OBCD joint pipeline

The development of joint phenotyping methods to investigate mutant mice demonstrates the power of unbiased functional genomics in osteoarthritis gene discovery. Open access availability of IMPC mouse lines provides a rich resource for investigation of disease mechanisms and the identification and testing of preventive or disease modifying drugs. Here we describe three further distinct applications of the phenotyping pipeline that demonstrate how the OBCD joint phenotype database can be leveraged to add value to studies of human

osteoarthritis and how the imaging techniques can be applied to address additional hypotheses.

Knockout mice for genes differentially expressed in human OA cartilage

Phenotyping genetically modified mouse models is a powerful method to functionally annotate potential disease susceptibility genes identified in human studies. We interrogated the OBCD database of joint phenotypes from randomly selected mouse lines to annotate the function of 409 genes differentially expressed in low- versus high-grade articular cartilage in osteoarthritis patients[51]. Phenotype data in 16-week-old mice were available from eight differentially expressed genes (*Unk*, *Josd1,* gasdermin E (*Gsdme*)*,* Rho GTPase activating protein 30 (*Arhgap30*)*, Ccdc6, Col4a2,* methyl-CpG binding domain protein 1 (*Mbd1*) and staufen double-stranded RNA binding protein 2 (*Stau2*)). Mutation of six (75%) differentially expressed genes in mice resulted in joint abnormalities whereas deletion of the other two (25%) (*Mbd1*, *Stau2*) had no effect (Supplementary Data 4 and 5). This compares with 14 out of 50 randomly selected lines (28%) with strong evidence for a functional role in osteoarthritis following prioritization in this study (P=0.01582, Fisher's exact test). This enrichment of signal supports interrogation of the OBCD phenotype database to accelerate functional gene discovery in osteoarthritis. Combining complementary mouse and human gene discovery approaches demonstrates synergy, especially as none of the genes identified have been associated with osteoarthritis in GWAS.

Unk$^{-/-}$ mice had decreased articular cartilage volume in the LTP and decreased maximum cartilage thickness with increased surface damage affecting the MTP (Figure 5, Supplementary Figure 6, Supplementary Data 4, 5 and 7). No osteophytes were detected in *Unk*$^{-/-}$ mice (Supplementary Data 8). Mean body weight was 2.49 SDs below the wild-type mean (Supplementary Data 4). UNK is an RNA binding zinc finger protein expressed in developing brain that controls neuronal morphology[52]. *UNK* is also expressed in the skeleton (Supplementary Data 7). *Unk*$^{-/-}$ mice display articular cartilage loss with early cartilage damage

indicating *Unk* is a protective osteoarthritis susceptibility gene. Downregulation of *UNK* at the RNA and protein levels in high-grade osteoarthritis cartilage in humans suggests a role in disease pathogenesis.

Josd1$^{-/-}$ mice had decreased median and maximum cartilage thickness affecting both joint compartments and increased articular cartilage surface damage affecting the lateral compartment (Figure 5, Supplementary Figure 6, Supplementary Data 4, 5 and 7). No osteophytes were detected in *Josd1*$^{-/-}$ mice (Supplementary Data 8). JOSD1 is a widely expressed deubiquitinating enzyme that may play a role in regulation of cell membrane dynamics[53]. It stabilizes SOCS1, an important negative regulator of cytokine signalling[54, 55], and *JOSD1* mRNA is upregulated in high-grade osteoarthritis cartilage (Figure 5), suggesting a protective role for JOSD1 in osteoarthritis.

Gsdme$^{-/-}$ mice had increased articular cartilage surface damage affecting the lateral compartment (Figure 5, Supplementary Figure 6, Supplementary Data 4, 5 and 7). No osteophytes were detected in *Gsdme*$^{-/-}$ mice (Supplementary Data 8). Gasdermin E is a member of a family of proteins that facilitate necrotic programmed cell death (pyroptosis) following cleavage by caspase-3[56]. Pyroptosis has recently been shown to promote knee osteoarthritis[57, 58], and upregulation of *Gsdme* in high-grade osteoarthritis cartilage (Figure 5), supports a role for gasdermin E in disease progression.

Arhgap30$^{-/-}$ mice had increased articular cartilage surface damage affecting the lateral compartment and a significant outlier phenotype following Mahalanobis analysis (Supplementary Data 4and 7). *Arhgap30* encodes a Rho GTPase implicated in cell proliferation, migration and invasion acting via inhibition of Wnt[59]. Its putative role and increased expression in high-grade osteoarthritis cartilage (Figure 5), suggests involvement of ARHGAP30 in cartilage repair mechanisms during osteoarthritis pathogenesis.

Ccdc6$^{-/-}$ mice had increased articular cartilage surface damage affecting the medial compartment (Supplementary Data 4and 7). *Ccdc6* is a cell cycle checkpoint regulator that facilitates cell survival[60]. *CCDC6* was associated with heel BMD in a large GWAS[19] and *Ccdc6*$^{-/-}$ mice have decreased bone strength (Supplementary Data 7). Concordant with upregulation of CCDC6 in high-grade osteoarthritic cartilage (Figure 5), *CCDC6* was differentially expressed in a meta-analysis of gene expression profiling in synovial tissue from osteoarthritis cases and controls[61]. Overall, *CCDC6* represents an osteoarthritis susceptibility gene, deletion of which may contribute to development of osteoarthritis by actions in both cartilage and the synovium.

Col4a2$^{-/-}$ mice had increased articular cartilage surface damage affecting the lateral compartment (Supplementary Data 4). The collagen type IV alpha 2 chain is a major component of vascular basement membrane[62], but is also present in skeletal basement membranes due to expression in bone cells (Supplementary Data 7). *COL4A2* was differentially expressed in osteoarthritis synovium[61], in knee joints following DMM surgery[63], in a time course analysis following DMM surgery[64], and in articular cartilage biopsies from osteoarthritis patients[65]. The *COL4A2* locus was differentially methylated in a genome-wide analysis of hip compared to knee osteoarthritis cartilage[66]. In summary, several lines of evidence suggest a role for *Col4a2* in the pathogenesis of osteoarthritis and the articular cartilage damage in *Col4a2*$^{-/-}$ mice is consistent with this hypothesis (Figure 1 and Supplementary Data 4).

Age-related joint degeneration

We next studied 4- and 12-month-old WT mice to investigate the effect of ageing. Joints from 12-month-old mice had features of osteoarthritis compared to 4-month-old juvenile mice. Even though articular volume and thickness did not change with age, the 12-month-old mice had significantly increased areas of cartilage damage in both lateral and medial compartments. Histology revealed an increased maximum OARSI score on the LTP in 12-month-old

compared with 4-month-old mice. These changes were accompanied by loss of subchondral bone (decreased SC BV/TV, SC Tb.Th and SC Tb.N) in the lateral compartment and increased SC BMC in the medial compartment (Figure 6 and Supplementary Figure 7).

Dio2^{Ala92} mice are protected from osteoarthritis

We next used the phenotyping pipeline to investigate a mouse model of a common human single nucleotide polymorphism previously associated with human osteoarthritis. As proof-of-concept for functional investigation of signals arising from human genetic association studies, we studied the effect of rs225014, a polymorphism in the human *DIO2* gene. rs225014 results in a substitution at amino acid 92 (Thr92Ala) in the DIO2 enzyme that activates thyroid hormones in target cells. The minor allele (Ala92) frequency is estimated at 40%[67]. rs225014 was positively associated with osteoarthritis in a genome-wide linkage study that included replication in a separate cohort[68], but this association was not reproduced in the Rotterdam study[69] or a later meta-analysis[70]. Nevertheless, a polymorphism in the *DIO2* promoter (rs12885300) has been associated with hip geometry in genome-wide linkage studies. While minor alleles at rs225014 and rs12885300 were associated with opposite osteoarthritis outcomes, a haplotype combining both was associated with increased disease susceptibility[71]. Overall, these findings implicate regulation of local thyroid hormone availability in the pathogenesis of osteoarthritis and are consistent with well-known actions of thyroid hormones, which stimulate hypertrophic chondrocyte differentiation[72] and expression of cartilage matrix degrading enzymes[73]. Nevertheless, *DIO2* has not been associated with osteoarthritis in GWAS, and the role of the *DIO2* rs225014 polymorphism remains controversial, attracting much debate even beyond the osteoarthritis field[74, 75, 76].

In vivo studies have demonstrated increased subchondral bone but intact articular cartilage in *Dio2$^{-/-}$* knockout mice[77]. *Dio2$^{-/-}$* mice have decreased calreticulin expression in articular cartilage[78], a gene implicated in cartilage thinning in response to mechanical loading[79], and are protected from cartilage damage following forced exercise. By contrast, increased *DIO2*

expression resulted in negative effects on chondrocyte function and homeostasis *in vitro*[80] and *DIO2* expression was increased in articular cartilage from osteoarthritis patients[81]. Furthermore, cartilage-specific overexpression of *Dio2* in transgenic rats resulted in cartilage destruction[82]. Together, these studies suggest that decreased DIO2 expression and reduced thyroid hormone availability in the joint may protect against osteoarthritis, whereas increased DIO2 expression may increase susceptibility.

To test this hypothesis, we analyzed the joint phenotypes of *Dio2*^{Thr92} and *Dio2*^{Ala92} mutant mice previously generated using the *CRISPR/Cas9* system[74]. The Thr92Ala polymorphism results in decreased enzyme activity and impaired conversion of the prohormone thyroxine (T4) to the active hormone triiodothyronine (T3) reducing local thyroid hormone signalling[74]. Analysis of joints from 16-week-old male mice demonstrated features of early onset osteoarthritis in *Dio2*^{Thr92} mice compared to *Dio2*^{Ala92} mice. *Dio2*^{Thr92} mice had decreased cartilage volume and median thickness with increased articular cartilage damage (Figure 7 and Supplementary Figure 8). By contrast, *Dio2*^{Ala92} mutants had no signs of osteoarthritis, indicating a protective role for the Ala92 polymorphism and providing the first functional evidence of a role for this candidate *DIO2* polymorphism *in vivo*. The data provide further evidence that decreased thyroid hormone signalling is protective against osteoarthritis.

Discussion

The molecular mechanisms that initiate and drive osteoarthritis progression and their genetic origins are largely unknown. We developed rapid-throughput joint phenotyping methods to identify abnormal joint phenotypes in mutant mice, identifying 14 genes with strong functional evidence for involvement in osteoarthritis pathogenesis. Previous approaches to mouse joint phenotyping have used modifications to standard μCT methods, including the use of cartilage staining or phase-contrast, while others have used confocal microscopy[83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94]. Here, we combine three complementary methods; iodine contrast-enhanced micro-computerized tomography (ICEμCT), joint surface replication (JSR) and subchondral bone X-ray microradiography (scXRM), in the first comprehensive multiplex mouse joint phenotyping pipeline. One important consideration regarding the highly sensitive JSR method is the very low level of cartilage damage that is present in wild type mice (0-2.8%). Thus, the method may be less reliable for quantitation of decreased cartilage damage in phenotype screening studies. Nevertheless, the method correlates well with OARSI histology and provides reliable and sensitive detection of increased cartilage damage, which is a pathognomonic feature of osteoarthritis. Furthermore, it is particularly useful in studies of mice with protective mutations that have been challenged by DMM provocation surgery providing such studies are adequately powered.

To demonstrate the broad utility of this joint phenotyping pipeline we describe four distinct applications (i) Screening of randomly selected KO lines (ii) Analysis of knockout mice with deletion of genes differentially expressed in human OA cartilage, (iii) Analysis of mice with age-related joint degeneration and (iv) Analysis of a mouse model of a common human single nucleotide polymorphism associated with osteoarthritis.

We first performed joint phenotyping in 50 randomly selected IMPC knockout lines generated by the Sanger Institute and identified one or more outlier parameters in 50% of lines. This is

similar to the frequency of outlier parameters identified during bone phenotyping of randomly-selected IMPC knockout mice[16, 17, 19] and consistent with the numerous independent loci associated with osteoarthritis[1, 2, 3] and osteoporosis[17, 19] in large scale GWAS. We, therefore, developed a robust and unbiased prioritisation pipeline that identified only those genes with multi-parameter joint abnormalities (n=14), which were further prioritised based on biological evidence, plausibility and relevance to human disease, finally highlighting 4 out of 50 genes (*Bhlhe40, Pitx1*, *Sh3bp4* and *Unk*).

Pitx1$^{+/-}$ mice had the most severe and extensive joint destruction in these studies, with severe erosions, complete loss of articular cartilage and osteophyte formation. Decreased *PITX1* mRNA expression in primary human articular chondrocytes, reduced PITX1 protein in histological sections of human osteoarthritic cartilage, and increased subchondral bone thickening in preliminary studies of ageing *Pitx1*$^{+/-}$ mice are all consistent with an important role for PITX1 in osteoarthritis pathogenesis[95]. Overall, we show that heterozygous deletion of *Pitx1,* a critical developmental gene, results in severe early onset joint damage. *Pitx1* is thus an osteoarthritis gene with a key role in disease onset and progression that likely involves its effects on joint morphology and the developmental programmes that are re-initiated during disease progression.

We exploited our joint phenotype database to characterize to functions of six candidate genes differentially expressed in human osteoarthritis cartilage**.** These studies identify roles for *Unk*, *Josd1*, *Gsdme*, *Arhgap30, Ccdc6* and *Col4a2* in the pathogenesis of osteoarthritis. *Unk*$^{-/-}$ mice were also prioritized as one of four lines with the strongest combined evidence for a key functional role in osteoarthritis pathogenesis, while *Josd1*$^{-/-}$ mice were one of seven lines with the most severely abnormal joint phenotype (Figure 1 and Supplementary Data 7). Overall, these findings demonstrate the value of leveraging the combined power of mouse and human experimental pipelines to enhance understanding of complex disease.

Previous histological studies have only demonstrated the onset of osteoarthritis at 15 months of age in wild-type *C57BL6* mice[96]. The current studies, however, clearly define early features of age-related joint degeneration in one-year old animals. Thus, our phenotyping methods are more economical and have improved sensitivity to detect early osteoarthritis compared to current labour- and resource-intensive histological techniques. OBCD joint phenotyping is one third of the cost and takes one sixth of the time compared to the current gold standard OARSI histological analysis (Supplementary Figure 9, Supplementary Data 9, 10 and 11). In addition to identifying genes that provoke early onset of osteoarthritis when deleted, these methods now provide efficient opportunities to analyze genes that protect joints from age-related degeneration when inhibited or deleted, or for analysis of drug intervention studies.

Finally, we identified a protective role for a minor allele of the *DIO2* gene (Ala92). The Ala92 variant impairs conversion of the prohormone T4 to the active hormone T3 and is controversially associated with osteoarthritis. Furthermore, it has been suggested that expression of this minor allele underlies the continuing psychological and metabolic symptoms experienced by some hypothyroid patients despite restoration of normal serum T4 levels following levothyroxine replacement[76, 97]. As a result, genetic testing and treatment of hypothyroid patients with the Ala92 variant using increased doses of levothyroxine (T4) or a combination of levothyroxine and liothyronine (T3) has been proposed. Importantly, this therapeutic approach, which aims to overcome impaired conversion of T4 to T3 by the Ala92 variant, may actually have detrimental long-term consequences as our data indicate that expression of the Ala92 variant may maintain cartilage integrity. Thus, the allelic imbalance that has been reported to increase expression of the Ala92 variant in human osteoarthritis cartilage[81], may in fact represent a counter-regulatory response in damaged articular cartilage to protect against osteoarthritis progression. Our findings underscore the importance of caution and further research in this controversial area, and have important public health implications, as levothyroxine is the most commonly prescribed drug in the USA and third most commonly prescribed in the UK[98].

In these studies, we performed DMM provocation surgery in 10-week-old mice as described originally by Glasson[13, 99]. Nevertheless, others now recommend undertaking DMM surgery after 12-weeks-of-age because of more advanced skeletal maturity[100]. We performed sham surgery in the contralateral knee in order to include paired control samples and minimise animal numbers in accordance with the principles of reduction, refinement and replacement (3Rs). Nevertheless, we acknowledge that sham surgery may influence joint mobility and affect phenotype severity[100]. This study identifies abnormal joint phenotypes in several knockout mice suggesting the deleted genes have important roles in joint development, maintenance and repair. However, such findings do not necessarily indicate that these genes represent tractable drug targets in the prevention and treatment of OA. Nevertheless, we also report that *CRISPR/Cas9* targeting of a single amino acid in the *Dio2* gene identifies a role for thyroid hormone signalling in joint homeostasis and has important clinical implications.

In summary, we have shown our joint phenotyping methods have broad applications in osteoarthritis research that will accelerate functional gene discovery, advance understanding of disease pathogenesis and identify drug targets for this debilitating chronic disease.

Acknowledgements

We thank Jayashree Bagchi-Chakraborty and Mahrokh Nodani for technical assistance and Carmen Huesa for assistance with synovitis scoring. We thank members of Sanger Mouse Pipelines (Mouse Informatics, Molecular Technologies, Genome Engineering Technologies, Mouse Production Team, Mouse Phenotyping) and the Research Support Facility for provision and management of mice. This work was funded by a Wellcome Trust Strategic Award (101123). Generation of mutant mice was funded by the Wellcome Trust (098051). G.R.W., J.H.D.B. were funded by a Wellcome Trust Joint Investigator Award (110140, 110141) and European Commission Horizon 2020 Grant (666869), E.Z. by the Wellcome Trust (206194), and PIC by Mrs. Janice Gibson and the Ernest Heine Family Foundation.

Author Contributions

AB, EZ, PIC, GRW and JHDB conceived and designed experiments, NCB, VDL, JGL, JAW, AB, GRW and JHDB developed methods, NCB, JGL and JHDB wrote software code, NCB, JS, RJ, GRW and JHDB performed statistical analyses, NCB, KFC, HD, DKE, NSM, AA, VDL, JGL, JAW, EG, LS and EAM conducted experiments, JS, LS, SEY, JMW, VEV, FK, JKW, DJA, CJL, ACB, PIC, GRW and JHDB provided experimental resources, VDL curated data, AB, EZ, PIC, GRW and JHDB acquired funding, NCB, GRW and JHDB produced figures, NCB, GRW and JHDB wrote the manuscript, and all authors reviewed and edited the manuscript.

Competing Interests Statement

The authors declare no competing interests

Figure Legends

Figure 1

Rapid-throughput joint phenotyping identifies osteoarthritis genes

Knee joints from randomly selected 16-week-old mutant mice generated at the Wellcome Trust Sanger Institute (WTSI) for the International Mouse Phenotyping Consortium (IMPC) are analyzed by three imaging modalities; iodine contrast-enhanced micro-computerized tomography (ICEμCT), joint surface replication (JSR) and subchondral bone X-ray microradiography (scXRM). The methods were validated by comparison with scoring of histological sections (Osteoarthritis Research Society International (OARSI) histology after destabilization of the medial meniscus (DMM) provocation surgery. 25 mouse lines (3-way Venn diagram) had outlier phenotypes relative to wild-type reference data following statistical analyses. Further prioritization was based on additional skeletal abnormalities identified in mutant mice, gene expression, association with human disease and literature searching. Boxes and 4-way Venn diagram show top ranked genes in each category. *Pitx1*, *Bhlhe40*, *Sh3bp4* and *Unk* were the top ranked genes. Green text: most severely abnormal joint phenotypes. Red boxes indicate mutant lines with increased articular cartilage surface damage. Source data are provided as a Source Data file.

Figure 2

Imaging methods

a. Coronal views of conventional μCT scan of a 16-week-old wild-type tibia (top left) and ICEμCT scan (top right), at 2μm voxel size; mineralized tissue is white. In ICEμCT scan contrast agent at similar X-ray absorption to mineralized tissue; soft tissue including articular cartilage is black. Coronal and sagittal views of ICEμCT scans with VOIs (volumes of interest) scaled according to a-b (medial-lateral) and c-d (anterior-posterior) tibial dimensions, and positioned midway between plateau edges (medial plateau, MTP, e-f, lateral plateau, LTP, g-h, with a 7% shift medially). Plateau midpoints: dashed lines. M: medial, L: lateral, A: anterior,

P: posterior. ICEμCT VOIs used for quantitative analyses (MTP, bottom left, LTP, bottom right); subchondral bone (blue); articular cartilage (red). Scale bars = 500μm.

b. Upper images: surface of disarticulated 16-week-old wild-type mouse tibia; limbs immersed in impression medium for moulding; mould (brown) overlaid with casting resin; surface of mould after removal of cast; surface of cast. Lower images: detail of mould surface; detail of cast surface; carbon-coated cast; back-scattered electron-scanning electron microscopy (BSE-SEM) image of whole tibia; high-power showing fibrillations in articular cartilage surface following DMM surgery.

c. Upper images: BSE-SEM view of lateral tibial plateau; selection of whole plateau region for analysis; automated identification of surface damage edges; thresholding of damage to capture all damage areas; damage detection. Lower images: high power views demonstrate damage detail. Scale bars = 500μm (upper), 100μm (lower).

d. Upper images: Faxitron MX20; 16-week-old wild-type mouse tibia alongside aluminium (left), plastic (right) and steel (bottom) standards; plain X-ray grey scale image of tibia and standards; greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Lower images: pseudocoloured image of proximal tibia with scaled (to width of tibia, white line) subchondral regions of interest (ROI), white boxes (LTP; lateral tibial plateau, MTP; medial tibial plateau); high power pseudocoloured views of LTP and MTP ROIs; grey scale pixel distribution in relation to standards; grey scale distribution stretched to plastic and steel standards. Scale bar = 1mm. mm; millimeter, mgHA/cm^3; milligrams of hydroxyapatite/cubic centimeter.

Figure 3

Severe early onset osteoarthritis in Pitx1$^{+/-}$ mice is similar to extent of joint damage following osteoarthritis provocation surgery

a. Iodine contrast enhanced μCT images of medial tibial plateau articular cartilage (red) and subchondral bone (blue) from 22-week-old wild type (WT) male mice 12 weeks after sham

operation or following destabilization of medial meniscus surgery (DMM; n=11) to provoke osteoarthritis, and 16-week-old male WT and heterozygous *Pitx1*$^{+/-}$ mice (n=7). Graphs show decreased articular cartilage volume (Cg.V) and maximum articular cartilage thickness (Max Cg.Th).

b. Back-scattered electron-scanning electron microscopy images of tibial plateau joint surface replicas from sham and DMM-operated WT mice, and WT and *Pitx1*$^{+/-}$ mice. Graphs show increased articular cartilage (Cg.) damage area.

c. X-ray microradiography images of proximal tibia and the medial tibial plateau subchondral bone region of interest (dashed box) from sham and DMM-operated WT mice, and WT and *Pitx1*$^{+/-}$ mice in which greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Graphs show increased subchondral BMC (SC BMC; DMM only).

d. Coronal sections of knee joint compartments stained with Safranin-O/Fast green from sham and DMM-operated WT mice, and from WT and two examples of *Pitx1*$^{+/-}$ mice with extensive joint destruction. Graph shows the maximum OARSI score on the medial tibial plateau in DMM-operated mice (n=11). Arrows indicate areas of cartilage destruction. Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters (SC BMC), reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters (Cg.V, Max Cg.Th, Cg. Damage Area), reference range is the 2.5th-97.5th percentile, and dashed line is the median. Scale bars = 100μm (a, b, d) and 1mm (c). DMM vs. Sham: **P*<0.05, ***P*<0.01, ****P*<0.001, 2-tailed Wilcoxon matched pairs signed rank test for Cg.V (*P*=0.00131), Max. Cg.Th (*P*=0.017334), Cg. Damage Area (*P*=0.000031) and Max. OARSI score (*P*=0.000977), and paired 2-tailed *t*-test for SC BMC (*P*=0.000477). *Pitx1*$^{+/-}$: **P*<0.00568, ****P*<0.0001, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. Cg.V: *P*=0.00008, Max. Cg.Th: *P*=0.00319, Cg. Damage area: *P*=0.00001. mm; millimeter. Source data are provided as a Source Data file.

Figure 4

Early onset osteoarthritis in Bhlhe40$^{-/-}$ and Sh3pb4$^{-/-}$ mice

a. Iodine contrast enhanced μCT images of lateral tibial plateau articular cartilage (red) and subchondral bone (blue) from 16-week-old wild type (WT) and homozygous *Bhlhe40$^{-/-}$* and *Sh3bp4$^{-/-}$* mice. Graphs show articular cartilage volume (Cg.V) and median articular cartilage thickness (Median Cg.Th) in *Bhlhe40$^{-/-}$* (n=3) and *Sh3bp4$^{-/-}$* (n=5) mice.

b. Back-scattered electron-scanning electron microscopy images of lateral tibial plateau joint surface replicas from WT, *Bhlhe40$^{-/-}$* and *Sh3bp4$^{-/-}$* mice. Graph shows articular cartilage (Cg) damage.

c. X-ray microradiography images of proximal tibia and the lateral tibial plateau subchondral bone region of interest (dashed box) from WT, *Bhlhe40$^{-/-}$* and *Sh3bp4$^{-/-}$* mice in which greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Graph shows no change in subchondral BMC (SC BMC).

d. Coronal sections of lateral knee joint compartment stained with Safranin-O/Fast green from WT, *Bhlhe40$^{-/-}$* and *Sh3bp4$^{-/-}$* mice (n=3 sections from 1 mouse each). Arrows indicate areas of cartilage damage. Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters (Cg.V, SC BMC), reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters (Median Cg.Th, Cg. Damage Area), reference range is the 2.5th-97.5th percentile, and dashed line is the median. Scale bars = 100μm (a, b, d) and 1mm (c). **P*<0.00568, ***P*<0.001, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. Cg.V: *P*=0.00311 (*Sh3bp4$^{-/-}$*), Median Cg.Th: *P*=0.00469 (*Bhlhe40$^{-/-}$*), *P*=0.00081 (*Sh3bp4$^{-/-}$*), Cg. Damage Area: *P*=0.00282 (*Sh3bp4$^{-/-}$*). mm; millimeter. Source data are provided as a Source Data file.

Figure 5

Early onset osteoarthritis in mice with deletion of genes differentially expressed in human osteoarthritis cartilage

a. 16-week-old mice with deletion of candidate genes differentially expressed in osteoarthritic human articular cartilage (*Unk, Josd1, Gsdme, Arhgap30, Ccdc6, Col4a2, Mdb1* and *Stau2*) were investigated for osteoarthritis phenotypes.

b. Differential expression (log fold change, logFC) of human orthologues of candidate genes in low-grade versus high-grade osteoarthritis in human articular cartilage samples. Genes with significantly increased expression in high-grade osteoarthritis cartilage shown in green and those with decreased expression in red.

c. Iodine contrast enhanced μCT images of lateral tibial plateau articular cartilage (red) and subchondral bone (blue) from wild type (WT), *Unk*$^{-/-}$, *Josd1*$^{-/-}$ and *Gsdme*$^{-/-}$ mice. Graphs show articular cartilage volume (Cg.V) and maximum articular cartilage thickness (Max Cg.Th) in *Unk*$^{-/-}$ (n=4), *Josd1*$^{-/-}$ (n=6) and *Gsdme*$^{-/-}$ (n=4) mice.

d. Back-scattered electron-scanning electron microscopy images of tibial plateau joint surface replicas from WT, *Unk*$^{-/-}$, *Josd1*$^{-/-}$ and *Gsdme*$^{-/-}$ mice. Graph shows articular cartilage (Cg) damage.

e. X-ray microradiography images of proximal tibia and the lateral tibial plateau subchondral bone region of interest (dashed box) from WT, *Unk*$^{-/-}$, *Josd1*$^{-/-}$ and *Gsdme*$^{-/-}$ mice in which greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Graph shows subchondral BMC (SC BMC Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters (Cg.V, Max. Cg.Th, SC BMC), reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters (Cg. Damage Area), reference range is the 2.5th-97.5th percentile, and dashed line is the median. Scale bars = 100μm (c, d) and 1mm (e). **P*<0.00568, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. Cg.V: *P*=0.00215 (*Unk*$^{-/-}$), Max. Cg.Th *P*=0.00362 (*Josd1*$^{-/-}$), Cg. Damage Area:

P=0.00563 (*Josd1*$^{-/-}$), *P*=0.00355 (*Gsdme*$^{-/-}$). mm; millimeter. Source data are provided as a Source Data file.

Figure 6

Age-related joint degeneration

a. Cohorts (n=6) of young adult (4 months) and aged (12 months) wild type (WT) mice were analyzed by iodine contrast enhanced μCT (ICEμCT), joint surface replication (JSR) and subchondral X-ray microradiography (scXRM).

b. ICEμCT images of lateral tibial plateau articular cartilage (red) and subchondral bone (blue). Graphs show median articular cartilage thickness (Median Cg.Th) and subchondral bone BV/TV (SC BV/TV).

c. Back-scattered electron-scanning electron microscopy images of lateral tibial plateau JSRs. Graph shows articular cartilage (Cg) damage.

d. ScXRM images of proximal tibia and the lateral tibial plateau subchondral bone region of interest (dashed box) in which greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Graph shows subchondral BMC (SC BMC).

e. Coronal sections of lateral knee joint compartment stained with Safranin-O/Fast green (*n*=6 mice/group). Graph shows maximum OARSI histological scores. Arrow indicates cartilage damage. Scale bars = 100μm (b, c, e) and 1mm (d). **P*<0.05, ***P*<0.01, 2-tailed *t*-test (SC BV/TV, *P*=0.0082) or 2-tailed Wilcoxon rank sum test (Cg. Damage area, *P*=0.0022 and Max. OARSI score, *P*=0.0240). mm; millimeter. Source data are provided as a Source Data file.

Figure 7

Dio2^{Ala92} mice are protected from osteoarthritis

a. 16-week-old mice with *Dio2* polymorphisms homologous to the Thr92Ala polymorphism in human *DIO2* (*Dio2^{Thr92}*, n=9 and *Dio2^{Ala92}*, n=10) were analyzed by iodine contrast enhanced

μCT (ICEμCT), joint surface replication (JSR) and subchondral X-ray microradiography (scXRM).

b. ICEμCT images of medial tibial plateau articular cartilage (red) and subchondral bone (blue). Graphs show articular cartilage volume (Cg.V) and median articular cartilage thickness (Median Cg.Th).

c. Back-scattered electron-scanning electron microscopy images of medial tibial plateau JSRs. Graph shows articular cartilage (Cg) damage.

d. ScXRM images of proximal tibia and the medial tibial plateau subchondral bone region of interest (dashed box) in which greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink. Graph shows subchondral BMC (SC BMC). Scale bars = 100μm (b, c) and 1mm (d). **P*<0.05, ***P*<0.01, 2-tailed Wilcoxon rank sum test. Cg.V: *P*=0.019, Median Cg.Th: *P*=0.006, Cg. Damage area: *P*=0.016. Source data are provided as a Source Data file.

Methods

Experimental models

Mouse lines generated in this study have been deposited in the European Mutant Mouse Archive (EMMA; https://www.infrafrontier.eu/) or are available on request. Animal experiments were performed and reported in accordance with ARRIVE guidelines[101]. All animal experiments were conducted under licence at Imperial College (project licence PPL70/8785) and the Wellcome Trust Sanger Institute (project licence P77453634 and PPL80/2485) in accordance with the Animals (Scientific Procedures) Act 1986 and recommendations of the Weatherall report. Animal experiments were approved by the Sanger or Imperial College Hammersmith Campus Animal Welfare Ethical Review Bodies (AWERB) as appropriate. Studies performed at the University of Chicago were approved by the Institutional Animal Care and Use Committee (IACUC) at Rush University Medical Center (16-077 and 15-033). Male mice were used for all joint phenotype analyses.

IMPC wild-type and mutant mouse strains

Samples from 16-week old male wild-type (WT) and genetically-modified mice designed with deletion alleles on the *C57BL/6N Taconic;C57BL/6N* background were generated as part of the Wellcome Trust Sanger Institute's (WTSI) Mouse Genetics Project, part of the International Mouse Phenotyping Consortium (IMPC; http://www.mousephenotype.org). Details on back-crossing status, health status, administered drug and procedures, husbandry and specific conditions (including housing, food, temperature and cage conditions) have been reported previously[102, 103] with the exception that temperature was 21±2°C, humidity was 55±10%, the diet used was Mouse Breeder Diet 5021 (Labdiet, London, UK) and the IMPReSS screen had 4 fewer tests (hair phenotyping, open field, hot plate and stress induced hypothermia). Further details available on request. Mouse strains, genotypes and Research Resource Identifiers (RRIDs) are included in Supplementary Data 12. Samples from male mice were used for the OBCD joint phenotyping pipeline and samples from females were used for the OBCD bone

phenotyping pipeline. The rationale that justified this approach was that (i) the gold-standard DMM surgical provocation model for osteoarthritis is 100% penetrant in male mice but has variable penetrance in females[24, 99], and (ii) the ovariectomy provocation model for post-menopausal osteoporosis can only be performed in females.

Wild-type male mice for provocation studies

Mice were housed at Imperial College London in individually ventilated cages (Techniplast, UK) on a 12h light/dark cycle at 20-23°C with humidity of 45-65%. Mice were housed in sibling groups and given access to chow diet and water *ad libitum*. Mice were weighed weekly and were within a range of 25-35g. Housing density was between 1-5 animals per cage (up to 25g) or between 1-4 animals per cage (over 25g). Health status was monitored by screening sentinel mice every 3 months, and health reports are available on request. No previous drugs, tests or procedures were administered except those specified, and no adverse events occurred. Mice were euthanized by exposure to increasing concentrations of CO_2 followed by cervical dislocation.

For surgical provocation of osteoarthritis, WT *C57BL/6* virgin male mice were purchased from Charles River Laboratories (Margate, Kent, UK; n=32). Mice had been maintained as homozygous WT and back-crossed more than 20 generations (https://www.criver.com/sites/default/files/Technical%20Resources/C57BL_6%20Mouse%20Model%20Information%20Sheet.pdf). Mice were housed 4 per cage in sibling groups with identical enrichment and allocated to surgical groups based on cage number. Weight was monitored before and for 48 hours after surgery. Order of surgery was based on cage number and numerical mouse ID and is available on request. Surgery was performed at 10 weeks of age.

For provocation of osteoarthritis by aging, WT *129/SV/C57BL/6J* male mice (16 weeks and 1 year of age, 12 total, 6 mice/group) were generated from a long-standing colony originally obtained from Jacques Samarut (École Normale Supérieure, Lyon, France)[104]. WT mice were maintained and back-crossed for more than 20 generations.

Dio2 mutant mice

Homozygous 16-week-old F3 male mice (19 mice total) in which the wild-type mouse deiodinase 2 (Dio2) allele at residue 92 (proline/proline) was replaced with human minor (alanine/alanine) or major (threonine/threonine) alleles by *CRISPR/Cas9* ($Dio2^{Ala92}$, n=9 and $Dio2^{Thr92}$, n=10) were previously generated by Applied Stemcell Inc. (Milpitas, CA, USA) and housed as described[74]. No previous drugs, tests or procedures were administered. Information regarding back-crossing status, weight, health status, husbandry and specific conditions (including housing, temperature, humidity and cage conditions) is available on request.

Human sample collection

The study design and conduct complied with all relevant regulations regarding the use of human study participants and were conducted in accordance to the criteria set by the Declaration of Helsinki. All patients provided written, informed consent prior to participation. Tissue samples were collected from 115 patients undergoing total joint replacement surgery in 4 cohorts: 12 knee osteoarthritis patients (cohort 1; 2 women, 10 men, age 50-88 years, mean 68 years); 20 knee osteoarthritis patients (cohort 2; 14 women, 6 men, age 54-82 years, mean 70 years); 13 hip osteoarthritis patients (cohort 3; 8 women, 5 men, age 44-84 years, mean 62 years); 70 knee osteoarthritis patients (cohort 4; 42 women, 28 men, age 38-84 years, mean 70 years). Matched low-grade and high-grade cartilage samples were collected from each patient. This work was approved by Oxford NHS REC C (10/H0606/20 and 15/SC/0132). Samples from knee osteoarthritis patients were collected under Human Tissue Authority license 12182, Sheffield Musculoskeletal Biobank, University of Sheffield, UK;

samples from hip osteoarthritis patients were collected under National Research Ethics approval reference 11/EE/0011, Cambridge Biomedical Research Centre Human Research Tissue Bank, Cambridge University Hospitals, UK.

Rapid-throughput OBCD joint phenotyping

Sample preparation

Hindlimbs from WT and mutant mice were skinned and fixed for 24-30h in 10% neutral buffered-formalin, rinsed twice and stored in 70% ethanol at 4°C. Samples were anonymized and randomly assigned to batches for rapid-throughput analysis. Prior to analysis, limbs were rehydrated in PBS + 0.02% sodium azide for >24h. Soft tissue was removed and the knee joint was disarticulated under a Leica MZ9 dissecting microscope (Leica Microsystems, UK) with the aid of fine forceps (Dumont #5, Cat#11252-20; Fine Science Tools, Germany) and 5mm spring scissors (Vannas Tubingen, Cat#15003-08, Fine Science Tools, Germany).

Iodine contrast-enhanced µCT

This method was developed to detect signs of osteoarthritis, including cartilage damage and subchondral bone sclerosis or bone loss. The joint is immersed in an iodinated contrast agent with similar X-ray absorption to bone and imaged (Figure 2). This allows segmentation of both the articular cartilage and underlying subchondral bone in the same volume of interest. The technique can also be used to analyze the condyles and head of the femur.

Rehydrated tibiae were blotted on low-linting tissue (Kimtech, UK) to remove all liquid. Single tibiae were placed into 6mm sample holders (Scanco Medical AG, Brüttisellen, Switzerland) in 180mg/mL ethiodized oil (Lipiodol Ultra, Guerbet Laboratories, Solihull, UK) mixed with sunflower oil. The tibial epiphysis (1-1.5mm) was scanned using a Scanco µCT-50 micro-computerized tomography (µCT) scanner with X-ray tube potential 70 kVp, 200 µA tube current, 2µm voxel size with 1500ms integration time and 1x averaging. Documented quality control scans were performed routinely as indicated by the manufacturer and conform to

reporting recommendations[105]. Images were processed using the Scanco software (μCT Tomography v6.4-2/Open VMS), with Gaussian filtration (standard deviation, sigma; 0.8, support; 1), and component labelling for the 20 (subchondral bone) and 5 (articular cartilage) largest particles. A 3 voxel (subchondral bone) and 5 voxel (articular cartilage) dilation/erosion was applied.

Analysis was performed using Xming 6.9.0.31 (©2005-2007 Colin Harrison) software implemented with the pUTTy client program (release 0.62 ©1997-2011 Simon Tatham). Tomographs were rotated twice to place the limb in a standard position. First, the tibia was rotated clockwise to point the anterior-most patellar surface vertically. Secondly, the tibia was rotated 90° to place the medial-lateral axis of the tibia facing the viewer (Figure 2a, coronal views). Medial and lateral tibial plateaux were analyzed separately. Equally sized medial and lateral VOIs were scaled automatically according to the size of the tibia (Figure 2a, medial-lateral dimensions, a-b; anterior-posterior dimensions, c-d). The width of each VOI was equal to 14% of the medial-lateral axis, which equated to 59% of the medial plateau width and 53% of the lateral plateau width. The anterior-posterior dimension of each VOI was 33% of the size of each tibia. VOIs were located at the same positions on each limb by determining the plateau edges (Figure 2a; e, f, g, h). Each VOI was then centered on the medial-lateral midpoint of the plateau (Figure 2a, dashed line), and the lateral plateau VOI was further offset by a distance equivalent to 7% of the medial-lateral axis (Figure 2a). The medial VOI midpoint was 61.5%, and the lateral VOI midpoint 68.5% (7% posterior shift) along the anterior-posterior axis. This accounted for the different shapes of the medial and lateral plateau, based on standard measurements determined from mean values obtained from 10 wild-type mice to capture the thickest part of each plateau (Supplementary Data 13).

For articular cartilage, all tissue below a density threshold of 200 mg hydroxyapatite/cubic centimeter (mg HA/cm^3) was analyzed. The volume and thickness (median and maximum) were determined. For subchondral bone, all tissue between the articular cartilage and the

growth plate above a density threshold of 300 mg HA/cm^3 was analyzed. Relative subchondral bone volume/tissue volume (BV/TV), tissue mineral density (TMD), trabecular thickness (Tb.Th) and trabecular number (Tb.N) were determined. Of note, the use of the Lipiodol contrast agent in ICEμCT results in a systematic and predictable increase in subchondral bone parameters (Supplementary Data 13). However, since all samples are processed and imaged identically this systematic increase does not compromise data interpretation as results represent relative comparisons.

For osteophytes, coronal sections located at the midpoint of the VOI were scored for the presence or absence of osteophytes by 3 blinded independent scorers (Supplementary Figure 3, Supplementary Data 2, 8).

To validate ICEμCT, we compared the maximum cartilage thickness (Cg.Th) determined from OARSI histological sections with the same parameter determined by ICEμCT. We measured Cg.Th in ICEμCT images and histological sections from 23 samples that had undergone OBCD rapid-throughput joint phenotyping and histological analysis (samples obtained from sham-operated (n=3) and DMM-operated (n=3) 12-weeks after surgery at 22-weeks-of-age, wild-type mice at 16- (n=7) and 52- (n=6) weeks-of-age, *Sh3bp4*$^{-/-}$ (n=1), *Pitx1*$^{-/-}$ (n=2), *Bhlhe40*$^{-/-}$ (n=1) at 16-weeks-of-age). Linear regression was performed to determine the correlation between the two methods and Bland-Altman plots were generated[83, 106]. Maximum Cg.Th determined by the two techniques was strongly correlated (R^2 = 0.91, MTP, P<0.0001, R^2 = 0.755, LTP, P<0.0001, Supplementary Figure 10). Furthermore, there was no proportional bias in measurements, and no increase in differences between methods with phenotype severity (Bland-Altman plot; MTP; bias, -0.0002273mm, R^2 = 0.125, P=0.106, LTP; bias, 0.004217mm, R^2 = 0.011, P=0.63, Supplementary Figure 10).

Nevertheless, the LTP Max Cg.Th was significantly thicker when measured by histology (mean value +4.4%) when compared to measurement by ICEμCT (*P=0.014) (Supplementary Figure 10e). To investigate this difference, we tested whether ICEμCT sample preparation

resulted in shrinkage. However, ICEμCT analysis of fresh and fixed samples showed no significant difference in cartilage volume and thickness (Supplementary Figure 10f). Overall, these studies demonstrate an absence of shrinkage during ICEμCT sample preparation, but a consistent difference in Max Cg.Th observed in the LTP but not the MTP when comparing ICEμCT with histology. These findings suggest a difference in anatomical orientation of the LTP relative to the MTP that is revealed when 2D (histology) and 3D (ICEμCT) imaging are compared. We hypothesise that the apparent increase in LTP thickness determined by histology results from a horizontal anterior-posterior tilt in the LTP relative to the MTP, but this relative anatomical relationship does not affect the Max Cg.Th measured in 3 dimensions by ICEμCT.

Joint surface replication

This method was developed to determine articular cartilage surface damage. The technique enables visualization of joint surface morphology at high-resolution in the natural hydrated condition by using a resin cast as a surrogate (Figure 2b-c). The method can also be used to analyze the condyles and head of the femur.

Moulding joint surface replicas

Disarticulated, rehydrated tibiae were positioned with the articular surfaces parallel to the moulding surface using Draper Helping Hand brackets with articulated arms (31324Hh, Bamford Trading, Ross-on-Wye, UK). A multi-aperture moulding template of depth >0.5cm was custom-made. Moisture was removed to preserve surface detail by blotting with low-linting tissue (Kimtech). Samples were immersed immediately in Virtual Light Body dental impression medium (Ivoclar-Vivadent, Leicester, UK) for 4 minutes prior to de-moulding and immediate return to storage medium. Moulds were visually inspected for defects or bubbles and either cast immediately or stored for a maximum of 2 weeks according to manufacturer's advice.

Casting joint surface replicas

Joint surface replicas were cast using Crystal Clear 202 acrylic resin (Smooth-On, Bentley Advanced Materials, London, UK) according to manufacturer's instructions (Figure 2). Pre-warmed (37°C) parts A and B (1:9 ratio) were mixed thoroughly and vacuum de-gassed. Resin was poured into joint surface moulds by gravity flow using a fine pipette tip with the aid of a dissecting microscope (MZ9, Leica Microsystems, UK) and bubbles were removed from mould-resin interfaces. Casts were set for 16h at 18-23°C, post-cured for 6h at 65°C and removed from moulds after 7 days at room temperature, according to manufacturer's instructions. Each cast was secured to a custom-produced 10x10x0.3cm aluminum raft with double-sided carbon tape (Agar Scientific, Stansted, UK). Rafts of samples were coated with >20nm carbon using a high-vacuum bench-top carbon coater (Agar Scientific).

Imaging joint surface replicas

Coated samples were analyzed using a Vega3 XMU scanning electron microscope (Tescan UK, Cambridge) at high vacuum with a 4-quadrant back-scattered electron detector (Deben, Bury St. Edmunds, UK). Tibial plateaux were imaged at 1536x1536 pixel image size with a scan speed of 5 and 3x frame-corrected averaging over a view field of 1800x1800μm (Figure 2). Beam voltage was 20kV and working distance was 17mm. ImageSnapper semi-automated imaging software (V1.0, Tescan UK) was used to acquire images. The presence or absence of osteophytes was determined by 3 blinded independent scorers (Supplementary Figure 3, Supplementary Data 2, 8).

Damage quantitation in joint surface replicas

A quantitation pipeline was developed to determine areas of damage on the joint surfaces of the medial and lateral plateau accurately, while excluding extraneous surface debris (Figure 2c). Macros (ImageJ1.44) were produced for automated analysis (see below; Supplementary

Data 14). Samples were analyzed in batches by a blinded operator. After each step, macros included batch-saving of modified images, unless otherwise stated. Particle sizes are based on pixel size of 1.37μm^2. Steps for semi-automated damage quantitation were:

1) Select plateau (*free select* tool; Figure 2c, 2nd panel and close-up). Isolated plateaux were batch-saved with a modified file name (Macro 1; automated)
2) Exclude bright outlier particles <4 pixels (5.5μm^2; Macro 2; automated)
3) Detect damage using the *find edges* tool (Macro 3, Figure 2c, 3rd panel and close-up*;* automated)
4) Manually erase debris (soft tissue, bubbles in moulded surfaces) with close comparison to the original image and batch save with a modified file name (Macro 4; automated)
5) Adjust the threshold to capture all damage accurately with close comparison to the original image and batch save with a modified file name (Macro 5, Figure 2c, 4th panel and close-up; automated)
6) Quantify damage area (*analyze particles* tool; Macro 6, Figure 2c, 5th panel and close-up) particle size >20 pixels (27.5μm^2), circularity 0-0.5 (0; line; 1; perfect circle). Excluding circular objects further refines the damage selection as circular objects are not likely to be genuine damage (automated)

Damage area was recorded and represented as percentage of the total plateau area. The whole plateau area was determined by performing steps 1–2 above, and then:

7) Set threshold to minimum and batch save with a modified file name (Macro 5; automated)
8) Determine whole plateau area *(analyze particles* tool; Macro 6; plateau outline in Figure 2c, 5th panel), particle size >100 pixels (188μm^2), circularity 0–1 (automated)

Subchondral X-ray microradiography

Quantitative subchondral X-ray microradiography (scXRM) was performed at 10μm pixel resolution using a Faxitron MX20 variable kV point-projection X-ray source and digital image

system (Qados, Cross Technologies plc, Berkshire, UK) with modifications to previously published protocols[21, 22]. The protocol is detailed in full below.

Disarticulated tibiae were imaged flat with the posterior surface facing downwards (Figure 2d). Samples were imaged at 26 kV for 15s with the sample tray raised 6mm above the 5x magnification slot to maximize analysis area. Quantitation was performed in ImageJ1.44 by stretching pixel information between density values for plastic (minimum) and steel (maximum) standards and assigning to one of 256 bins using Macro 7 and automated batch saving (Supplementary Data 14). The width of each tibia at the growth plate was determined and used to automatically scale the regions of interest (ROIs; height: 9%, width: 34%, of the growth plate width). These dimensions were based on an optimal ROI size determined from 10 wild-type tibiae that was sufficiently wide to include the whole subchondral plate and shallow enough to avoid calcified cartilage. The ROI positions were parallel to each plateau and included the topmost yellow pixel representing the bone directly beneath the tibial plateau. For each ROI, the number of pixels in each bin was determined using the *Custom Histogram* macro (http://rsb.info.nih.gov/ij/). The median grey level (density bin at which the cumulative frequency of pixels reaches 50%) defines the relative bone mineral content (BMC). BMC determined by scXRM correlated with BMC determined by ICEµCT (tissue mineral density multiplied by bone volume (TMD*BV) in 100 WT mice; LTP (R^2=0.601, ***P=2.77x10^{-11}) and MTP (R^2=0.531, ***P=2.77x10^{-11}; Supplementary Data 6).

Repeatability and precision

To determine the repeatability of each method, 7 wild-type samples were selected that were evenly distributed across the reference range. Samples were blinded and analyzed 5 times in a random order, and in a different random order for each method. Mean, standard error of the mean, absolute precision error [standard deviation; PE(SD)], and precision error as percentage of coefficient of variation [PE(%CV)][107] were calculated for each sample across 5

measurements (Supplementary Figure 1b and Supplementary Data 15). Results were also expressed as the number of standard deviations away from the reference range mean (Supplementary Data 15). The average precision error for each parameter was calculated using the root-mean-square[107] (Supplementary Data 15).

To determine whether the precision error of the methods increases with phenotype severity in mutants, we performed an additional repeatability analysis on a subset of 6 knockout mouse lines with a wide range of phenotype severities (*Bhlhe40*$^{-/-}$, *Gsdme*$^{-/-}$, *Josd1*$^{-/-}$, *Pitx1*$^{+-/-}$, *Sh3bp4*$^{-/-}$ and *Unk*$^{-/-}$). Samples were analysed in an identical manner to the wild-type cohort described above. Spearman correlations were calculated comparing phenotype severity with absolute precision error [PE(SD)] and with precision error as a percentage coefficient of variation [PE(%CV)]. Linear regression analysis demonstrated that precision error did not change with phenotype severity for 17/18 parameters. Nevertheless, the absolute precision error of MTP cartilage surface damage [Cg. Damage Area (%)] did increase with phenotype severity (R^2=0.906, *P*=0.001; Supplementary Data 16).

To assess repeatability of the methods in both wild-type and mutant cohorts, two-way mixed-model intra-class correlation coefficients (single measure, absolute agreement) with 95% confidence intervals were calculated for repeated analyses by a single rater[83, 108]. All parameters had intra-class correlation coefficients >0.8, indicating excellent reproducibility (Supplementary Data 15).

Surgical provocation of osteoarthritis

Mice underwent destabilization of the medial meniscus (DMM)[13, 24] osteoarthritis provocation surgery at 10 weeks-of-age. To minimize recovery time, mice were anaesthetized using inhaled isofluorane/oxygen mix (1-1.5L/min), and breathing rate was monitored visually throughout surgery. The medial meniscal-tibial ligament (MMTL) of the right knee was

transected using a 5mm microsurgical blade (World Precision Instruments, Hitchin, UK, Cat# 500249). Sham surgery was performed on the left knee, in which skin and joint capsule were opened but the MMTL was left intact. Each joint capsule was closed with a single suture, and the skin was closed using intradermal sutures. Temperature was maintained using a heat-pad and monitored by rectal thermometer. Post-operative buprenorphine (0.1mg/kg in saline, Vetergesic; Ceva United Kingdom, Amersham, UK), and carprofen (5mg/kg in saline, Rimadyl; Zoetis, Leatherhead, UK) were administered subcutaneously post-surgery while under anesthetic, and daily for 48h. Dosage was calculated based on individual weight. Physical pain indicators (including mobility, piloerection and hunching) were monitored post-surgery and daily for 48h. Surgeries were performed between 10am and 4pm in a designated surgical suite. Mice were returned to the home cage for recovery. Mice were housed 4 per cage in sibling groups and sacrificed 12 weeks post-surgery.

Left and right hindlimbs from half of the experimental cohort (16 mice; 2 mice/cage) were phenotyped by rapid-throughput OBCD joint phenotyping. The limbs of the other half of the surgical cohort (2 mice/cage) were sectioned and scored for osteoarthritis using gold standard protocols. Numbers of mice in each group were based on group numbers in published guidelines[24].

Histological scoring of DMM surgical samples

Articulated left (sham) and right (DMM) hindlimbs were decalcified for 7-10 days in 10% EDTA and embedded in paraffin wax in a 90° flexed physiological position. Coronal 4µm sections were cut at 80µm intervals through the knee joint and stained with Safranin O/Fast green and Weigert's haematoxylin counterstain[77, 109]. Sections from 5 levels within the articulating zone centered around the midpoint were scored for each limb. All sections were scored by 3 blinded observers using the Osteoarthritis Research Society International (OARSI) scoring system[23]. Summed and maximum scores were calculated for each compartment (medial tibial plateau;

MTP, medial femoral condyle; MFC, lateral tibial plateau; LTP, lateral femoral condyle; LFC) in a single limb from 5 sections. In cases when 5 sections were not available, scores were weighted according to the following equation:

$$\text{Weighted score} = \text{score} + [(5-N) \times (\text{score}/N)] \quad (1)$$

where *N* was the number of scoreable sections. Weighted scores were rounded to the nearest 0.5. Sum scores represent the sum of scores from 5 levels. The maximum score is the highest score (out of a maximum of 6). The total joint sum score is the sum of scores for all compartments across 5 levels. The total joint maximum score represents the maximum score from all compartments across 5 levels. The presence of osteophytes was also scored (Supplementary Data 1). Lack of displacement of the medial meniscus was considered a criterion for exclusion (n=1 mouse). Synovitis severity was scored in the same sections [(severity score = sum of pannus formation score (0-3) + synovial lining hyperplasia score (0-3) + sub-synovial inflammation score (0-3)][110, 111, 112]. Five sections from each knee were scored by a blinded scorer. Differences between mean scores were evaluated by 2-tailed Wilcoxon matched-pairs signed rank tests (Supplementary Figure 3, Supplementary Data 17).

Analysis of osteophytes

We scored osteophytes by histology, ICEμCT and JSR in sham and DMM-operated limbs 12 weeks after surgery at 22 weeks-of-age, and by ICEμCT and JSR in mutant mouse lines at 16 weeks-of-age. In sham and DMM-operated limbs, the numbers of samples in which osteophytes were detected by ICEμCT and JSR did not differ from the number of samples in which osteophytes were detected by histology ($P>0.9999$; $P=0.7874$, Fisher's exact test; Supplementary Data 2), indicating detection of osteophytes by ICEμCT and JSR is equivalent to detection by OARSI histology. A combination of ICEμCT and JSR is required for osteophyte

analysis in mutant mice, as JSR may not detect osteophytes in dysmorphic or severely damaged joints (Supplementary Data 8).

Co-registration with gold standard histology

To co-register data obtained from OBCD joint phenotype analysis directly with OARSI histological scoring, we performed further histology and OARSI scoring of samples from 3 mice that had previously been phenotyped in the OBCD pipeline and found to have mild, intermediate and severe degrees of joint damage (Supplementary Figure 4). 4µm coronal sections were cut at 50µm intervals through the disarticulated tibiae and femora. Five sections centered around the midpoint of each femur and tibia were analysed by 3 independent scorers, and the mean summed and maximum scores determined.

Functional genomics analysis of human cartilage

We isolated chondrocytes from each cartilage sample as described[113] and detailed in full in Supplementary Methods (see Supplementary Information). Assays and analyses were performed as described in Supplementary Methods (see Supplementary Information), and RNA sequencing was performed on the Illumina HiSeq2000 or Hiseq4000 (75bp paired-ends), with quality control including FastQC 0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). For raw RNA sequencing data details see Data Availability Statement. For protein extracts, we carried out digestion, 6-plex or 10-plex tandem mass tag labelling and peptide fractionation. For samples from 12 knee osteoarthritis patients, we applied a liquid chromatography mass spectrometry (LC-MS) analysis using the Dionex Ultimate 3000 ultra-high-performance liquid chromatography (UHPLC) system coupled with the high-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific GmbH, Dreieich, Germany). For all remaining samples, LC-MS analysis was performed on the Dionex Ultimate 3000 UHPLC system coupled with the

Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fisher Scientific). For proteomics data details see Data Availability Statement.

OBCD bone phenotyping

Samples from 16-week-old female IMPC knockout mice were used to determine bone mass and strength parameters in the rapid-throughput OBCD bone phenotyping pipeline [16, 17, 19], and results are reported in Supplementary Data 7. Samples were stored in 70% ethanol, anonymized and randomly assigned to batches for rapid-throughput analysis. The relative BMC and length of the femur and caudal vertebrae were determined at 10μm pixel resolution by digital X-ray microradiography (Faxitron MX20). Micro-CT (Scanco μCT-50, 70kV, 200μA, 0.5mm aluminum filter) was used to determine cortical bone parameters (thickness, BMD, internal diameter) at 10μm voxel resolution in a 1.5mm region centered on the mid-shaft region 56% of the way along the length of the femur distal to the femoral head, and trabecular bone parameters (bone volume, trabecular number, thickness, spacing) at 5μm voxel resolution in a 1mm region beginning 100μm proximal to the distal growth plate. Biomechanical variables of bone strength and toughness (yield load, maximum load, fracture load, energy dissipated prior to fracture) were derived from destructive three-point bend testing of the femur and compression testing of caudal vertebrae 6 and 7 (Instron 5543 load frame, 100N and 500N load cells). Overall, 19 skeletal parameters were reported for each individual mouse studied and compared to reference data obtained from 320 16-week-old wild type female mice of identical genetic background.

Prioritization pipeline

A prioritization pipeline was developed based on the severity of the mutant mouse phenotype and i) additional skeletal abnormalities identified in mutant mice, ii) gene expression in skeletal cells and tissues, iii) gene association with human monogenic and complex diseases, and iv) structured literature searching (Supplementary Data 7).

Genes were allocated overall priority scores out of a total maximum of 21 as follows:

A maximum score of 6 for rapid-throughput OBCD joint phenotyping (outlier according to criteria of reference range, Wilcoxon test followed by Bonferroni adjustment and Mahalanobis analysis, 1 point each; abnormal articular cartilage morphology, articular cartilage damage, subchondral bone morphology, 1 point each).

A maximum score of 5 for abnormal skeletal phenotypes reported in genetically modified mice by other studies. The following resources were searched: 1) OBCD bone phenotype data, 2) International Mouse Phenotyping Resource of Standardized Screens (IMPReSS; https://www.mousephenotype.org/impress), 3) Deciphering the Mechanisms of Developmental Disorders database (DMDD; https://dmdd.org.uk/), 4) PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) using the search criteria "GENE NAME AND (knockout OR deletion OR mutation)", and 5) Mouse Genome Informatics database (http://www.informatics.jax.org/). The presence of an abnormal skeletal phenotype was assigned 1 point per database.

A maximum score of 4 for expression in skeletal tissues and cells (the skeleton, chondrocytes, osteoblasts/osteocytes and osteoclasts; Supplementary Data 7). Expression in the skeleton (1 point) was determined by searching MGI (http://www.informatics.jax.org/) and BioGPS (http://biogps.org/#goto=welcome). An expression level greater than the median was considered to indicate tissue or cell type expression. Expression in chondrocytes (1 point) was determined by searching published transcriptome data[27] and SkeletalVis[28] (http://phenome.manchester.ac.uk/). Search was limited by species (human) in both datasets. In Kean et al.[27] a baseMean expression greater than the median was considered to indicate cell type expression. In SkeletalVis searches, a fold-change greater than 2 (P<0.001) was considered to indicate cell type expression. Expression in osteoblasts and/or osteocytes (1

point) was determined by searching BioGPS, SkeletalVis, and osteocyte RNAseq data[19]. Expression in osteoclasts (1 point) was determined by searching BioGPS, human osteoclast RNAseq datasets, and SkeletalVis. A baseMean expression greater than the median was considered to indicate cell type expression.

A maximum score of 4 for association with monogenic and polygenic skeletal disease (Supplementary Data 7). Monogenic disease association was determined by searching MGI for human-mouse disease association, and Online Mendelian Inheritance in Man (OMIM, https://www.omim.org/) by gene name (1 point). Polygenic disease association was assessed by searching the European Bioinformatics Institute Genome-wide Association Studies database (EBI GWAS Catalogue, https://www.ebi.ac.uk/gwas/) for association with arthritis and skeletal diseases (1 point each).

A maximum score of 2 for publications related to skeletal tissues and cells (Supplementary Data 7). Systematic structured searches of PubMed and Google Scholar (https://scholar.google.co.uk/) were performed using the search string “GENE NAME AND (arthritis OR cartilage OR skeleton OR chondrocyte OR osteoblast OR osteocyte OR osteoclast)”. For PubMed results, the following scores were assigned: 0 (<1 publication), 0.5 (1–24 publications), or 1 (>25 publications). For Google Scholar results, the following scores were assigned: 0 (<100 results), 0.5 (100-999 results), or 1 (>1000 results).

Quantitation and statistical analysis

Statistical details and results of experiments are included in the relevant results section, figure legends, Reporting Summary and below. For all experiments, n refers to biological replicates (number of mice or histological sections). Measurements were taken from distinct samples unless otherwise indicated. Exact *P*-values are reported in figure legends and Supplementary Data files.

OBCD joint phenotyping

Reference range data was calculated from 100 wild-type mice for each parameter, and the mean, standard deviation, median and percentiles determined (Supplementary Data 3). Frequency distribution of datasets was assessed by the Shapiro-Wilk normality test. Reference ranges 2 standard deviation above or below the mean (normally-distributed parameters), or between the 2.5^{th}-97.5^{th} percentiles (non-normally distributed parameters), were determined. An outlier was defined if the mean value of a phenotype parameter was outside the defined reference range.

Normally distributed parameters were articular cartilage volume (LTP), maximum articular cartilage thickness (LTP), subchondral bone volume/tissue volume (LTP), subchondral trabecular thickness (LTP, MTP), subchondral trabecular number (LTP), subchondral tissue mineral density (LTP), and subchondral bone mineral content (LTP, MTP). Non-normally distributed parameters were articular cartilage volume (MTP), median articular cartilage thickness (LTP, MTP), maximum articular cartilage thickness (MTP), articular cartilage surface damage (LTP, MTP), subchondral bone volume/tissue volume (MTP), subchondral trabecular number (MTP) and subchondral tissue mineral density (MTP).

To determine the statistical significance of outlier parameters identified relative to the reference range, we performed non-parametric 2-tailed Wilcoxon rank sum tests (Supplementary Data 5, *wilcox.test* function in R). For each parameter, this analysis calculated whether the distribution parameters in mutant samples was significantly different to the distribution of parameters in the 100-sample WT reference range. To account for multiple testing when defining statistical significance, we calculated the effective number of tests and then applied a Bonferroni correction as follows. Because several of the 18 phenotype parameters determined in the 100 WT mice used to calculate reference data were correlated

(Supplementary Data 6, *P<0.05), we first calculated the number of effective tests using WT data in R. We calculated the pairwise correlations between all 18 parameters and obtained eigenvalues of the correlation matrix using the *eigen* function in R. We then calculated the number of effective tests (N_{eff}) as:

$$N_{eff} = N - \sum_{\lambda} I(\lambda > 1) * (\lambda - 1) \quad (2)$$

where N=18 is the number of parameters, and λ denotes the eigenvalues, resulting in an effective N_{eff}≤=8.8. Thus, applying a Bonferroni correction for the effective number of tests, statistical significance for the differences between WT reference data and results from each mutant mouse line was P<0.00568, with a second more stringent threshold of P<0.0001 also correcting for the number of mouse lines. For each of 18 parameters across the 50 mouse lines, differences between mutant and WT values were determined by 2-tailed Wilcoxon rank sum tests.

Differences between young and old wild-type mice, and between *Dio2*^{Thr92} and *Dio2*^{Ala92} mice, were determined using unpaired 2-tailed *t*-tests or 2-tailed Wilcoxon rank sum tests.

Differences between sham and DMM-operated limbs from the same mouse were determined using paired 2-tailed *t*-tests, or 2-tailed Wilcoxon matched-pairs signed rank tests.

Differences between OARSI histology and synovitis scores in sections from sham and DMM-operated knee joints were determined using paired 2-tailed Wilcoxon matched-pairs signed rank tests. Differences between OARSI histology scores of sections from young and aged wild-type mice were determined using unpaired 2-tailed Wilcoxon rank sum tests.

Power calculations

To determine the sample size required, we determined coefficients of variation (CV, standard deviation/mean) for normally distributed parameters, and percentage median absolute deviation from the median (median absolute deviation from the median (MAD)/median) for non-normally distributed parameters (Supplementary Data 3). The effect size (d) was set at twice the standard deviation (σ) or half of the 95% range as a percentage of the median. To detect this effect size with 80% power, the number of mice per group (N) was determined by

$$N = 2\frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times \sigma^2}{d^2} \quad (3)$$

where power index $(Z_{1-\alpha/2}+Z_{1-\beta})^2$ is 7.85 for a power of 80%, with MAD replacing σ for non-normally distributed parameters. Thus, 4 mice are required to detect an effect size of 2 standard deviations at 80% power, and 6 mice at 95% power, for all parameters. For non-parametric data, 1-4 mice are required, depending on the parameter, to detect an outlier phenotype lying outside the 95% confidence intervals of the reference range (Supplementary Data 3). Minimum detectable effect sizes for all parameters at 80%, 90% and 95% power are presented in Supplementary Data 18.

Mahalanobis analysis

Robust Mahalanobis distances[114, 115] were determined to identify outliers in multivariate data. These distances measure how far each observation is from the center of a data cluster, taking into account the variances of the variables and the covariances of pairs of variables. Robust Mahalanobis distances (MD_i)

$$\mathrm{MD_i} = \sqrt{(x_i - T(X))C(X)^{-1}(x_i - T(X))^T} \quad (4)$$

were calculated for each sample (represented by a vector of multivariate observations x_i). T(X) is a robust (i.e. relative unaffected by outliers) estimate of the mean vector and C(X) is a

robust estimate of the covariance matrix. Under the assumption of multivariate normality, the distribution of MD_i^2 is approximately chi-squared with p degrees of freedom (where p is the number of variables). This means that any observation x_i with $MD_i^2 > \chi^2_{p;\,0.975}$ can be considered an outlier. The assumption of multivariate normality was checked by plotting the ordered distances against the corresponding quantiles of an appropriate chi-squared distribution. Robust estimates of the mean and covariance matrix are used so that potential outliers are not masked. The masking effect, by which outliers do not necessarily have a large Mahalanobis distance, can be caused by a cluster of outliers that attract the mean and inflate the covariance in its direct. By using a robust estimate of the sample mean and covariance, the influence of these outliers is removed and the Mahalanobis distance is able to expose all outliers. The minimum volume ellipsoid method was used to calculate robust estimates of the mean and covariance matrix. Given *n* observations and *p* variables, the minimum volume ellipsoid method seeks an ellipsoid containing

$$h = \left[\frac{(n+p+1)}{2}\right] \tag{5}$$

h points of minimum volume. All multivariate analysis was conducted in R (R Project for Statistical Computing). Overall, a mouse line was defined to have an abnormal joint phenotype using this method only when 50% or more individual samples from that line were identified as outliers following Mahalanobis analysis.

Analysis of RNA and protein in human tissue

Full details of this analysis are described in the Supplementary Methods (see Supplementary Information). Briefly, gene-level RNA quantification was carried out using Salmon 0.8.2[116] and tximport[117]. After quality control, we retained expression estimates for 15,249 genes with counts per million of 1 or higher in at least 40 cartilage samples (combining low-grade and

high-grade), and matched low-grade and high-grade cartilage samples from 83 patients. To identify gene expression differences between high-grade and low-grade cartilage, we carried out multiple analyses using limma[118, 119], DESeq2[120] and edgeR[121] with 5 different designs to account for potential confounders. This yielded 2557 genes with significant differential expression between low-grade and high-grade cartilage at 5% false discovery rate (FDR) across all 5 analysis designs and all 5 testing methods.

To carry out protein identification and quantification, we submitted the mass spectra to SequestHT search in Proteome Discoverer 2.1 and searched all spectra against a UniProt fasta file that contained 20,165 reviewed human entries. After quality control, we retained 4801 proteins that were quantified in at least 30% of samples, and matched low-grade and high-grade cartilage samples from 99 patients. To account for protein loading, abundance values were normalized by the sum of all protein abundances in a given sample, then log2-transformed and quantile normalised. An analysis for differential abundance was carried out using limma[118] with 2016 proteins showing significant differences at 5% FDR.

We used Ensembl38p10 to identify human orthologues for 50 mouse genes phenotyped by the OBCD joint phenotyping pipeline. We identified high-confidence one-to-one orthologues for 43/50 genes.

Data Availability Statement

Source data are provided with this paper. All datasets generated and/or analyzed during the current study are included in Supplementary and Source Data, and available from the corresponding authors on reasonable request or as detailed below. Figures 1, 3-7, and Supplementary Figures 1-8 and 10 have associated raw data (included in the Supplementary Data and Source Data). Additionally, raw RNA sequencing data have been deposited in the European Genome-Phenome Archive [EGA; https://ega-archive.org/) with the identifiers

EGAS00001002255, EGAD00001003355 (n=17), EGAD00001003354 (n=9), EGAD00001001331 (n=12). Proteomics data have been deposited in the PRoteomics IDEntifications database (PRIDE; https://www.ebi.ac.uk/pride/archive/) with the identifiers PXD006673, PXD00202014 and PDX014666, username: reviewer90654@ebi.ac.uk, password: AoeRBm3e). Public databases used in this study are: BioGPS (http://biogps.org/#goto=welcome), Deciphering the Mechanisms of Developmental Disorders database (DMDD; https://dmdd.org.uk/), European Mutant Mouse Archive (EMMA; https://www.infrafrontier.eu/), European Bioinformatics Institute Genome-wide Association Studies database (EBI GWAS Catalogue, https://www.ebi.ac.uk/gwas/), Google Scholar (https://scholar.google.co.uk/), International Mouse Phenotyping Resource of Standardized Screens (IMPReSS; https://www.mousephenotype.org/impress), Mouse Genome Informatics database (http://www.informatics.jax.org/), Online Mendelian Inheritance in Man (OMIM, https://www.omim.org/), PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) and SkeletalVis[28] (http://phenome.manchester.ac.uk/). There are no restrictions on data availability.

Code availability

Code generated during this project is provided as Supplementary Data 14 and in the public repository Github (https://github.com/Molendo/OBCD). Repository: Molendo. Project: OBCD (search term: Molendo/OBCD).

References

1. Styrkarsdottir U*, et al.* Meta-analysis of Icelandic and UK data sets identifies missense variants in SMO, IL11, COL11A1 and 13 more new loci associated with osteoarthritis. *Nat Genet* **50**, 1681-1687 (2018).

2. Tachmazidou I*, et al.* Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nat Genet* **51**, 230-236 (2019).

3. Zengini E*, et al.* Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nat Genet* **50**, 549-558 (2018).

4. Lories RJ, Luyten FP. The bone-cartilage unit in osteoarthritis. *Nature reviews Rheumatology* **7**, 43-49 (2011).

5. Choi WS*, et al.* The CH25H-CYP7B1-RORalpha axis of cholesterol metabolism regulates osteoarthritis. *Nature* **566**, 254-258 (2019).

6. Zhang M*, et al.* Induced superficial chondrocyte death reduces catabolic cartilage damage in murine posttraumatic osteoarthritis. *J Clin Invest* **126**, 2893-2902 (2016).

7. Dreier R. Hypertrophic differentiation of chondrocytes in osteoarthritis: the developmental aspect of degenerative joint disorders. *Arthritis Res Ther* **12**, 216 (2010).

8. Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* **64**, 1697-1707 (2012).

9. Burr DB, Gallant MA. Bone remodelling in osteoarthritis. *Nature Rev Rheumatol* **8**, 665-673 (2012).

10. Kuhn K, D'Lima DD, Hashimoto S, Lotz M. Cell death in cartilage. *Osteoarthritis Cartilage* **12**, 1-16 (2004).

11. Enomoto H*, et al.* Vascular endothelial growth factor isoforms and their receptors are expressed in human osteoarthritic cartilage. *Am J Pathol* **162**, 171-181 (2003).

12. Billinghurst RC*, et al.* Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. *J Clin Invest* **99**, 1534-1545 (1997).

13. Glasson SS*, et al.* Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature* **434**, 644-648 (2005).

14. Boyden LM*, et al.* High bone density due to a mutation in LDL-receptor-related protein 5. *N Engl J Med* **346**, 1513-1521 (2002).

15. Farooqi IS*, et al.* Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N Engl J Med* **341**, 879-884 (1999).

16. Bassett JH*, et al.* Rapid-throughput skeletal phenotyping of 100 knockout mice identifies 9 new genes that determine bone strength. *PLoS Genet* **8**, e1002858 (2012).

17. Kemp JP*, et al.* Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat Genet* **49**, 1468-1475 (2017).

18. Medina-Gomez C*, et al.* Life-Course Genome-wide Association Study Meta-analysis of Total Body BMD and Assessment of Age-Specific Effects. *Am J Hum Genet* **102**, 88-102 (2018).

19. Morris JA*, et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* **51**, 258-266 (2019).

20. Trajanoska K*, et al.* Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *BMJ* **362**, k3225 (2018).

21. Butterfield NC, Logan JG, Waung J, Williams GR, Bassett JHD. Quantitative X-Ray Imaging of Mouse Bone by Faxitron. *Methods Mol Biol* **1914**, 559-569 (2019).

22. Waung JA*, et al.* Quantitative X-ray microradiography for high-throughput phenotyping of osteoarthritis in mice. *Osteoarthritis Cartilage*, (2014).

23. Glasson SS, Chambers MG, Van Den Berg WB, Little CB. The OARSI histopathology initiative - recommendations for histological assessments of osteoarthritis in the mouse. *Osteoarthritis Cartilage* **18 Suppl 3**, S17-23 (2010).

24. Glasson SS, Blanchet TJ, Morris EA. The surgical destabilization of the medial meniscus (DMM) model of osteoarthritis in the 129/SvEv mouse. *Osteoarthritis Cartilage* **15**, 1061-1069 (2007).

25. Perez-Garcia V*, et al.* Placentation defects are highly prevalent in embryonic lethal mouse mutants. *Nature* **555**, 463-468 (2018).

26. Wu C, Jin X, Tsueng G, Afrasiabi C, Su AI. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res* **44**, D313-316 (2016).

27. Kean TJ, Ge Z, Li Y, Chen R, Dennis JE. Transcriptome-Wide Analysis of Human Chondrocyte Expansion on Synoviocyte Matrix. *Cells* **8**, 85 (2019).

28. Soul J, Hardingham T, Boot-Handford R, Schwartz JM. SkeletalVis: An exploration and meta-analysis data portal of cross-species skeletal transcriptomics data. *Bioinformatics* **35**, 2283–2290 (2019).

29. Szeto DP*, et al.* Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. *Genes Dev* **13**, 484-494 (1999).

30. Lanctot C, Moreau A, Chamberland M, Tremblay ML, Drouin J. Hindlimb patterning and mandible development require the Ptx1 gene. *Development* **126**, 1805-1810 (1999).

31. Spielmann M*, et al.* Homeotic arm-to-leg transformation associated with genomic rearrangements at the PITX1 locus. *Am J Hum Genet* **91**, 629-635 (2012).

32. Gurnett CA*, et al.* Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. *Am J Hum Genet* **83**, 616-622 (2008).

33. Boudjelal M, Taneja R, Matsubara S, Bouillet P, Dolle P, Chambon P. Overexpression of Stra13, a novel retinoic acid-inducible gene of the basic helix-loop-helix family, inhibits mesodermal and promotes neuronal differentiation of P19 cells. *Genes Dev* **11**, 2052-2065 (1997).

34. MacLean HE, Kronenberg HM. Expression of Stra13 during mouse endochondral bone development. *Gene Expr Patterns* **4**, 633-636 (2004).

35. Miyazaki K, Kawamoto T, Tanimoto K, Nishiyama M, Honda H, Kato Y. Identification of functional hypoxia response elements in the promoter region of the DEC1 and DEC2 genes. *J Biol Chem* **277**, 47014-47021 (2002).

36. Shen M*, et al.* Basic helix-loop-helix protein DEC1 promotes chondrocyte differentiation at the early and terminal stages. *J Biol Chem* **277**, 50112-50120 (2002).

37. Iwata T*, et al.* Effects of overexpression of basic helix-loop-helix transcription factor Dec1 on osteogenic and adipogenic differentiation of mesenchymal stem cells. *Eur J Cell Biol* **85**, 423-431 (2006).

38. Hu S*, et al.* Differentiated embryonic chondrocytes 1 expression of periodontal ligament tissue and gingival tissue in the patients with chronic periodontitis. *Arch Oral Biol* **60**, 517-525 (2015).

39. Wu L, Huang X, Li L, Huang H, Xu R, Luyten W. Insights on biology and pathology of HIF-1alpha/-2alpha, TGFbeta/BMP, Wnt/beta-catenin, and NF-kappaB pathways in osteoarthritis. *Curr Pharm Des* **18**, 3293-3312 (2012).

40. Saito T, Kawaguchi H. HIF-2alpha as a possible therapeutic target of osteoarthritis. *Osteoarthritis Cartilage* **18**, 1552-1556 (2010).

41. Gamer LW*, et al.* The Role of Bmp2 in the Maturation and Maintenance of the Murine Knee Joint. *J Bone Miner Res* **33**, 1708-1717 (2018).

42. Dai MW*, et al.* Parathyroid hormone(1-34) exhibits more comprehensive effects than celecoxib in cartilage metabolism and maintaining subchondral bone micro-architecture in meniscectomized guinea pigs. *Osteoarthritis Cartilage* **24**, 1103-1112 (2016).

43. Tosoni D*, et al.* TTP specifically regulates the internalization of the transferrin receptor. *Cell* **123**, 875-888 (2005).

44. Francavilla C*, et al.* Functional proteomics defines the molecular switch underlying FGF receptor trafficking and cellular outputs. *Mol Cell* **51**, 707-722 (2013).

45. Kim YM*, et al.* SH3BP4 is a negative regulator of amino acid-Rag GTPase-mTORC1 signaling. *Mol Cell* **46**, 833-846 (2012).

46. Antas P*, et al.* SH3BP4 Regulates Intestinal Stem Cells and Tumorigenesis by Modulating beta-Catenin Nuclear Localization. *Cell Rep* **26**, 2266-2273 e2264 (2019).

47. Rauner M*, et al.* Transferrin receptor 2 controls bone mass and pathological bone formation via BMP and Wnt signaling. *Nat Metab* **1**, 111-124 (2019).

48. Tang J*, et al.* Fibroblast Growth Factor Receptor 3 Inhibits Osteoarthritis Progression in the Knee Joints of Adult Mice. *Arthritis Rheum* **68**, 2432-2443 (2016).

49. Zhang Y*, et al.* Cartilage-specific deletion of mTOR upregulates autophagy and protects mice from osteoarthritis. *Ann Rheum Dis* **74**, 1432-1440 (2015).

50. Monteagudo S, Lories RJ. Cushioning the cartilage: a canonical Wnt restricting matter. *Nat Rev Rheum* **13**, 670-681 (2017).

51. Steinberg J*, et al.* Decoding the genomic basis of osteoarthritis. *bioRxiv*, doi: 10.1101/835850 (2019).

52. Murn J*, et al.* Control of a neuronal morphology program by an RNA-binding zinc finger protein, Unkempt. *Genes Dev* **29**, 501-512 (2015).

53. Seki T, Gong L, Williams AJ, Sakai N, Todi SV, Paulson HL. JosD1, a membrane-targeted deubiquitinating enzyme, is activated by ubiquitination and regulates membrane dynamics, cell motility, and endocytosis. *J Biol Chem* **288**, 17145-17155 (2013).

54. Wang X*, et al.* JOSD1 Negatively Regulates Type-I Interferon Antiviral Activity by Deubiquitinating and Stabilizing SOCS1. *Viral Immunol* **30**, 342-349 (2017).

55. Kapoor M, Martel-Pelletier J, Lajeunesse D, Pelletier JP, Fahmi H. Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nat Rev Rheum* **7**, 33-42 (2011).

56. Rogers C, Erkes DA, Nardone A, Aplin AE, Fernandes-Alnemri T, Alnemri ES. Gasdermin pores permeabilize mitochondria to augment caspase-3 activation during apoptosis and inflammasome activation. *Nat Commun* **10**, 1689 (2019).

57. Zhang L*, et al.* Increased HIF-1alpha in Knee Osteoarthritis Aggravate Synovial Fibrosis via Fibroblast-Like Synoviocyte Pyroptosis. *Oxid Med Cell Longev* **2019**, 6326517 (2019).

58. Zhao LR*, et al.* NLRP1 and NLRP3 inflammasomes mediate LPS/ATPinduced pyroptosis in knee osteoarthritis. *Mol Med Rep* **17**, 5463-5469 (2018).

59. Mao X, Tong J. ARHGAP30 suppressed lung cancer cell proliferation, migration, and invasion through inhibition of the Wnt/beta-catenin signaling pathway. *Onco Targets Ther* **11**, 7447-7457 (2018).

60. Thanasopoulou A, Stravopodis DJ, Dimas KS, Schwaller J, Anastasiadou E. Loss of CCDC6 affects cell cycle through impaired intra-S-phase checkpoint control. *PLoS One* **7**, e31007 (2012).

61. Fei Q*, et al.* Identification of upstream regulators for synovial expression signature genes in osteoarthritis. *Joint Bone Spine* **83**, 545-551 (2016).

62. Khoshnoodi J, Pedchenko V, Hudson BG. Mammalian collagen IV. *Microsc Res Tech* **71**, 357-370 (2008).

63. Yang X, Zhao J, He Y, Huangfu X. Screening for characteristic genes in osteoarthritis induced by destabilization of the medial meniscus utilizing bioinformatics approach. *J Musculoskelet Neuronal Interact* **14**, 343-348 (2014).

64. Olex AL, Turkett WH, Fetrow JS, Loeser RF. Integration of gene expression data with network-based analysis to identify signaling and metabolic pathways regulated during the development of osteoarthritis. *Gene* **542**, 38-45 (2014).

65. Karlsson C*, et al.* Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. *Osteoarthritis Cartilage* **18**, 581-592 (2010).

66. Rushton MD*, et al.* Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. *Arthritis Rheum* **66**, 2450-2460 (2014).

67. Peeters RP*, et al.* Polymorphisms in thyroid hormone pathway genes are associated with plasma TSH and iodothyronine levels in healthy subjects. *J Clin Endocrinol Metab* **88**, 2880-2888 (2003).

68. Meulenbelt I*, et al.* Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* **17**, 1867-1875 (2008).

69. Kerkhof HJ*, et al.* A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* **62**, 499-510 (2010).

70. Evangelou E*, et al.* A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. *Ann Rheum Dis* **73**, 2130-2136 (2014).

71. Waarsing JH*, et al.* Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. *Arthritis Rheum* **63**, 1349-1354 (2011).

72. Robson H, Siebler T, Stevens DA, Shalet SM, Williams GR. Thyroid hormone acts directly on growth plate chondrocytes to promote hypertrophic differentiation and inhibit clonal expansion and cell proliferation. *Endocrinology* **141**, 3887-3897 (2000).

73. Makihira S*, et al.* Thyroid hormone enhances aggrecanase-2/ADAM-TS5 expression and proteoglycan degradation in growth plate cartilage. *Endocrinology* **144**, 2480-2488 (2003).

74. Jo S*, et al.* Type 2 deiodinase polymorphism causes ER stress and hypothyroidism in the brain. *J Clin Invest* **129**, 230-245 (2019).

75. Hernandez A. Cognitive function in hypothyroidism: what is that deiodinase again? *J Clin Invest* **129**, 55-57 (2019).

76. Gereben B, McAninch EA, Ribeiro MO, Bianco AC. Scope and limitations of iodothyronine deiodinases in hypothyroidism. *Nat Rev Endocrinol* **11**, 642-652 (2015).

77. Waung JA, Bassett JH, Williams GR. Adult mice lacking the type 2 iodothyronine deiodinase have increased subchondral bone but normal articular cartilage. *Thyroid* **25**, 269-277 (2015).

78. Bomer N*, et al.* Aberrant Calreticulin Expression in Articular Cartilage of Dio2 Deficient Mice. *PLoS One* **11**, e0154999 (2016).

79. Li H*, et al.* Endoplasmic reticulum stress regulates rat mandibular cartilage thinning under compressive mechanical stress. *J Biol Chem* **288**, 18172-18183 (2013).

80. Bomer N*, et al.* Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. *Ann Rheum Dis* **74**, 1571-1579 (2015).

81. Bos SD*, et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Ann Rheum Dis* **71**, 1254-1258 (2012).

82. Nagase H*, et al.* Deiodinase 2 upregulation demonstrated in osteoarthritis patients cartilage causes cartilage destruction in tissue-specific transgenic rats. *Osteoarthritis Cartilage* **21**, 514-523 (2013).

83. Stok KS*, et al.* Three-Dimensional Quantitative Morphometric Analysis (QMA) for In Situ Joint and Tissue Assessment of Osteoarthritis in a Preclinical Rabbit Disease Model. *PLoS One* **11**, e0147564 (2016).

84. Lee YS*, et al.* Articular cartilage imaging by the use of phase-contrast tomography in a collagen-induced arthritis mouse model. *Acad Radiol* **17**, 244-250 (2010).

85. Moodie JP, Stok KS, Muller R, Vincent TL, Shefelbine SJ. Multimodal imaging demonstrates concomitant changes in bone and cartilage after destabilisation of the medial meniscus and increased joint laxity. *Osteoarthritis Cartilage* **19**, 163-170 (2011).

86. Kotwal N, Li J, Sandy J, Plaas A, Sumner DR. Initial application of EPIC-muCT to assess mouse articular cartilage morphology and composition: effects of aging and treadmill running. *Osteoarthritis Cartilage* **20**, 887-895 (2012).

87. Kerckhofs G, Sainz J, Wevers M, Van de Putte T, Schrooten J. Contrast-enhanced nanofocus computed tomography images the cartilage subtissue architecture in three dimensions. *Eur Cells Mater* **25**, 179-189 (2013).

88. Ruan MZ, Dawson B, Jiang MM, Gannon F, Heggeness M, Lee BH. Quantitative imaging of murine osteoarthritic cartilage by phase-contrast micro-computed tomography. *Arthritis Rheum* **65**, 388-396 (2013).

89. Das Neves Borges P, Forte AE, Vincent TL, Dini D, Marenzana M. Rapid, automated imaging of mouse articular cartilage by microCT for early detection of osteoarthritis and finite element modelling of joint mechanics. *Osteoarthritis Cartilage* **22**, 1419-1428 (2014).

90. Lakin BA*, et al.* Contrast-enhanced CT using a cationic contrast agent enables non-destructive assessment of the biochemical and biomechanical properties of mouse tibial plateau cartilage. *J Orthop Res* **34**, 1130-1138 (2016).

91. Mashiatulla M*, et al.* Murine articular cartilage morphology and compositional quantification with high resolution cationic contrast-enhanced muCT. *J Orthop Res* **35**, 2740-2748 (2017).

92. Staines KA, Madi K, Javaheri B, Lee PD, Pitsillides AA. A Computed Microtomography Method for Understanding Epiphyseal Growth Plate Fusion. *Front Mater* **4**, 48 (2018).

93. de Bournonville S, Vangrunderbeeck S, Kerckhofs G. Contrast-Enhanced MicroCT for Virtual 3D Anatomical Pathology of Biological Tissues: A Literature Review. *Contrast Media Mol Imaging* **2019**, 8617406 (2019).

94. Gabner S, Bock P, Fink D, Glosmann M, Handschuh S. The visible skeleton 2.0: phenotyping of cartilage and bone in fixed vertebrate embryos and foetuses based on X-ray microCT. *Development* **147**, doi:10.1242/dev.187633 (2020).

95. Picard C, Azeddine B, Moldovan F, Martel-Pelletier J, Moreau A. New emerging role of pitx1 transcription factor in osteoarthritis pathogenesis. *Clin Orthop Relat Res* **462**, 59-66 (2007).

96. Wilhelmi G, Faust R. Suitability of the C57 black mouse as an experimental animal for the study of skeletal changes due to ageing, with special reference to osteo-arthrosis and its response to tribenoside. *Pharmacology* **14**, 289-296 (1976).

97. Bianco AC, Kim BS. Pathophysiological relevance of deiodinase polymorphism. *Curr Opin Endocrinol Diabetes Obes* **25**, 341-346 (2018).

98. Rodriguez-Gutierrez R, Maraka S, Ospina NS, Montori VM, Brito JP. Levothyroxine overuse: time for an about face? *Lancet Diabetes Endocrinol* **5**, 246-248 (2017).

99. Ma HL, Blanchet TJ, Peluso D, Hopkins B, Morris EA, Glasson SS. Osteoarthritis severity is sex dependent in a surgical mouse model. *Osteoarthritis Cartilage* **15**, 695-700 (2007).

100. Fang H, Beier F. Mouse models of osteoarthritis: modelling risk factors and assessing outcomes. *Nat Rev Rheumatol* **10**, 413-421 (2014).

101. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* **8**, e1000412 (2010).

102. Skarnes WC*, et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337-342 (2011).

103. White JK*, et al.* Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**, 452-464 (2013).

104. Gauthier K*, et al.* Genetic analysis reveals different functions for the products of the thyroid hormone receptor alpha locus. *Mol Cell Biol* **21**, 4748-4760 (2001).

105. Bouxsein ML, Boyd SK, Christiansen BA, Guldberg RE, Jepsen KJ, Muller R. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J Bone Miner Res* **25**, 1468-1486 (2010).

106. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* **25**, 141-151 (2015).

107. Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* **5**, 262-270 (1995).

108. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* **86**, 420-428 (1979).

109. Schmitz N, Laverty S, Kraus VB, Aigner T. Basic methods in histopathology of joint tissues. *Osteoarthritis Cartilage* **18 Suppl 3**, S113-116 (2010).

110. Huesa C*, et al.* Proteinase-activated receptor 2 modulates OA-related pain, cartilage and bone pathology. *Ann Rheum Dis* **75**, 1989-1997 (2016).

111. Sophocleous A, Huesa C. Osteoarthritis Mouse Model of Destabilization of the Medial Meniscus. *Methods Molecular Biol* **1914**, 281-293 (2019).

112. Jackson MT*, et al.* Depletion of protease-activated receptor 2 but not protease-activated receptor 1 may confer protection against osteoarthritis in mice through extracartilaginous mechanisms. *Arthritis Rheum* **66**, 3337-3348 (2014).

113. Steinberg J*, et al.* Widespread epigenomic, transcriptomic and proteomic differences between hip osteophytic and articular chondrocytes in osteoarthritis. *Rheumatology* **57**, 1481-1489 (2018).

114. Mitchell AFS, Krzanowski WJ. The Mahalanobis distance and elliptic distributions. *Biometrika* **72**, 464-467 (1985).

115. Rousseeuw PJ, Van Zomeren BC. Unmasking Multivariate Outliers and Leverage Points. *J Am Stat Assoc* **85**, 633-639 (1990).

116. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).

117. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).

118. Ritchie ME*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

119. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).

120. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

121. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**, 4288-4297 (2012).

122. Steinberg J*, et al.* Integrative epigenomics, transcriptomics and proteomics of patient chondrocytes reveal genes and pathways involved in osteoarthritis. *Sci Rep* **7**, 8935 (2017).

Figure 1

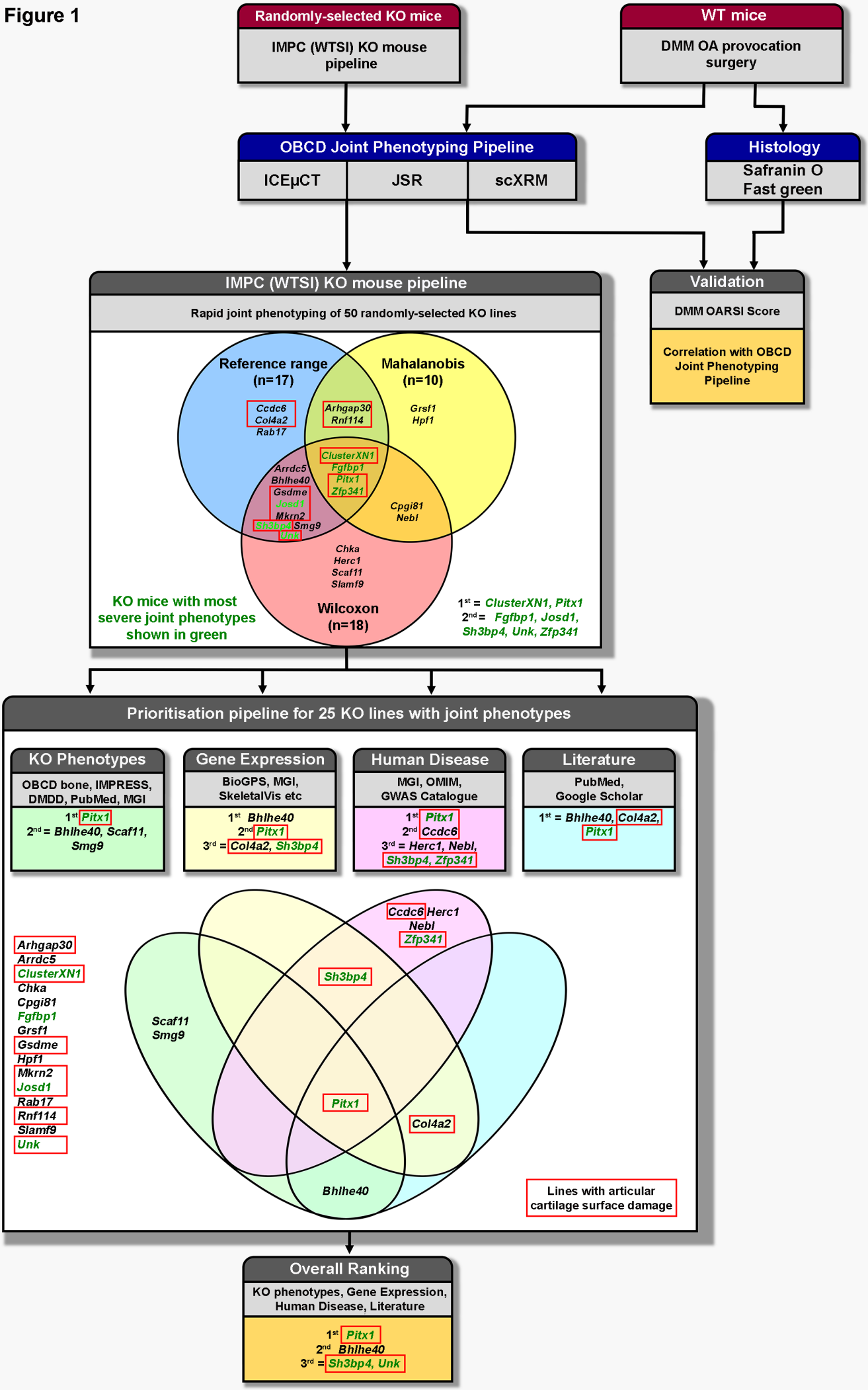

Randomly-selected KO mice
IMPC (WTSI) KO mouse pipeline
WT mice
DMM OA provocation surgery
OBCD Joint Phenotyping Pipeline
ICEμCT
JSR
scXRM
Histology
Safranin O
Fast green
IMPC (WTSI) KO mouse pipeline
Rapid joint phenotyping of 50 randomly-selected KO lines
Reference range (n=17)
Ccdc6
Col4a2
Rab17
Mahalanobis (n=10)
Arhgap30
Rnf114
Grsf1
Hpf1
ClusterXN1
Fgfbp1
Pitx1
Zfp341
Arrdc5
Bhlhe40
Gsdme
Josd1
Mkrn2
Sh3bp4
Smg9
Unk
Cpgi81
Nebl
Chka
Herc1
Scaf11
Slamf9
Wilcoxon (n=18)
KO mice with most severe joint phenotypes shown in green
1st = ClusterXN1, Pitx1
2nd = Fgfbp1, Josd1, Sh3bp4, Unk, Zfp341
Validation
DMM OARSI Score
Correlation with OBCD Joint Phenotyping Pipeline
Prioritisation pipeline for 25 KO lines with joint phenotypes
KO Phenotypes
OBCD bone, IMPRESS, DMDD, PubMed, MGI
1st Pitx1
2nd = Bhlhe40, Scaf11, Smg9
Gene Expression
BioGPS, MGI, SkeletalVis etc
1st Bhlhe40
2nd Pitx1
3rd = Col4a2, Sh3bp4
Human Disease
MGI, OMIM, GWAS Catalogue
1st Pitx1
2nd Ccdc6
3rd = Herc1, Nebl, Sh3bp4, Zfp341
Literature
PubMed, Google Scholar
1st = Bhlhe40, Col4a2, Pitx1
Arhgap30
Arrdc5
ClusterXN1
Chka
Cpgi81
Fgfbp1
Grsf1
Gsdme
Hpf1
Mkrn2
Josd1
Rab17
Rnf114
Slamf9
Unk
Ccdc6 Herc1 Nebl Zfp341
Sh3bp4
Scaf11 Smg9
Pitx1
Col4a2
Bhlhe40
Lines with articular cartilage surface damage
Overall Ranking
KO phenotypes, Gene Expression, Human Disease, Literature
1st Pitx1
2nd Bhlhe40
3rd = Sh3bp4, Unk

Figure 2

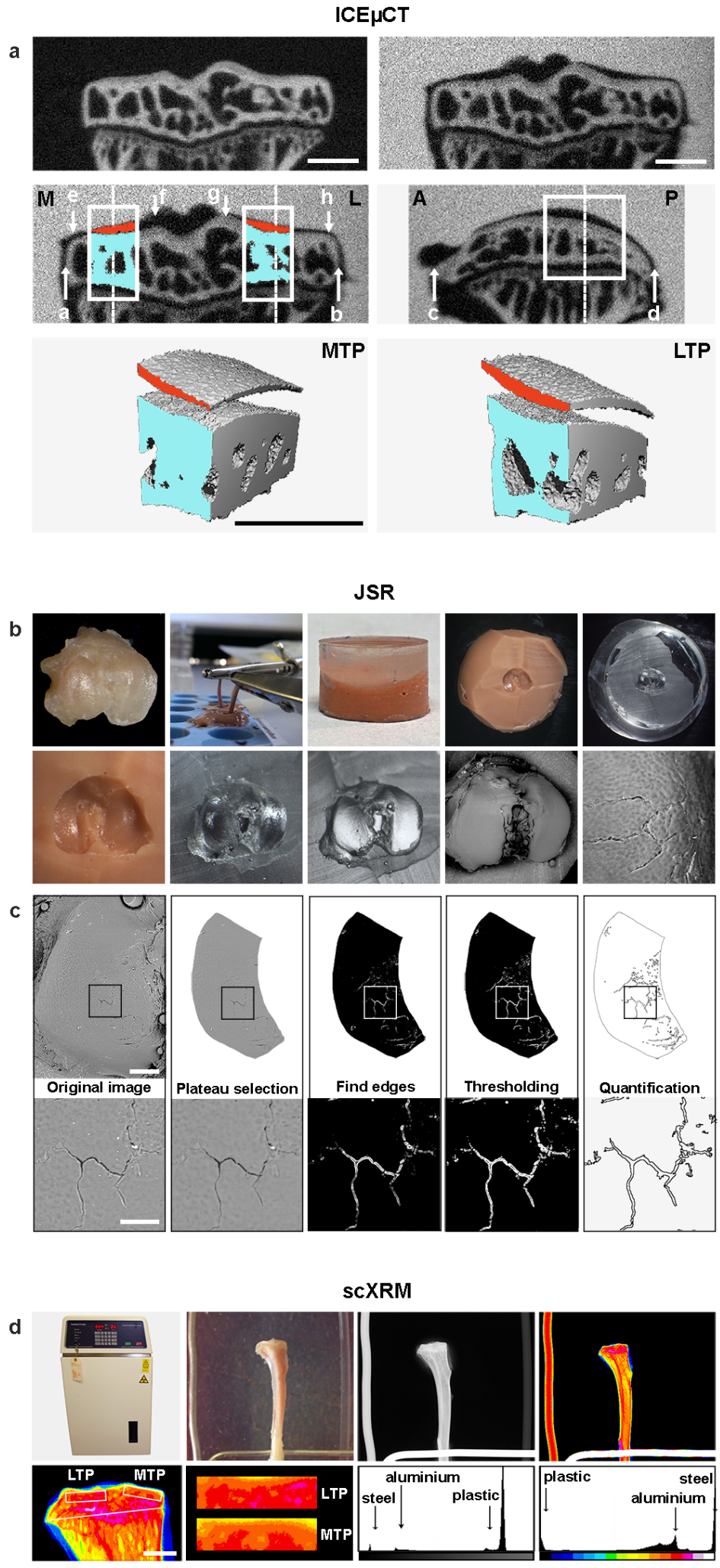

ICEµCT
a
M
e
f
g
h
L
a
b
A
P
c
d
MTP
LTP
JSR
b
c
Original image
Plateau selection
Find edges
Thresholding
Quantification
scXRM
d
LTP
MTP
LTP
MTP
aluminium
steel
plastic
plastic
steel
aluminium

Figure 3

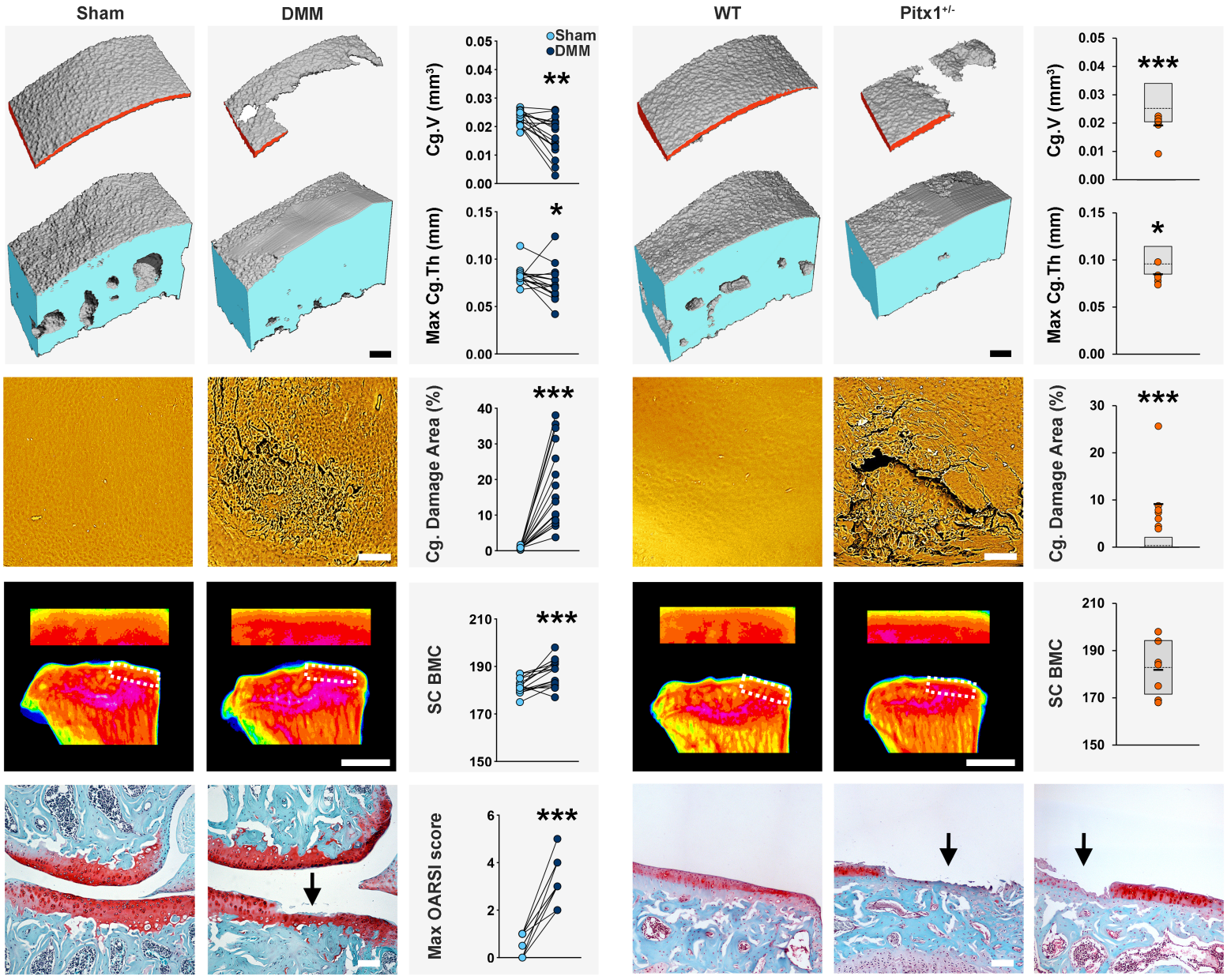

Sham
DMM
WT
Pitx1+/-
a
b
c
d
Sham
DMM
Cg.V (mm³)
Max Cg.Th (mm)
Cg. Damage Area (%)
SC BMC
Max OARSI score
**
*
***

Figure 4

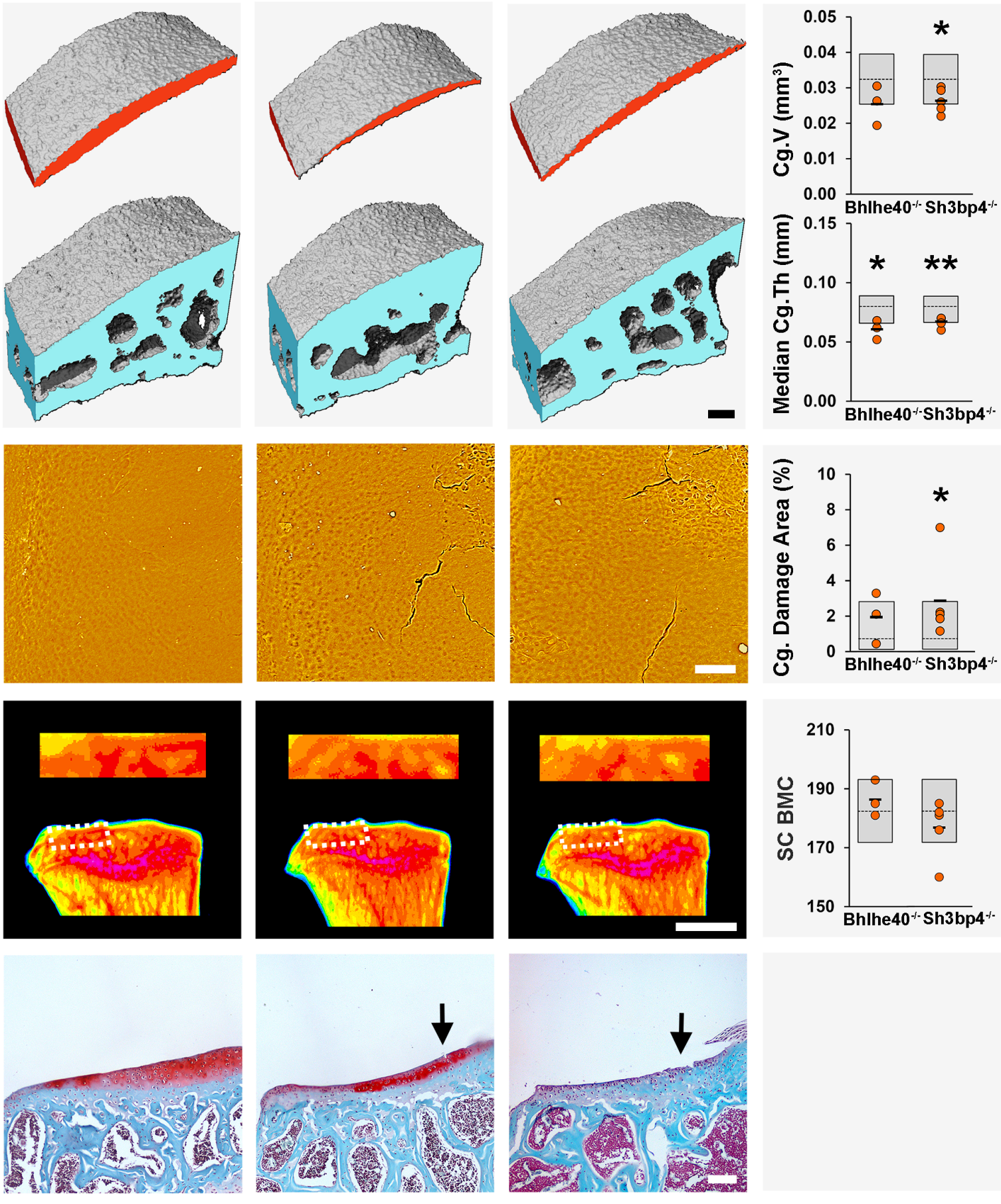
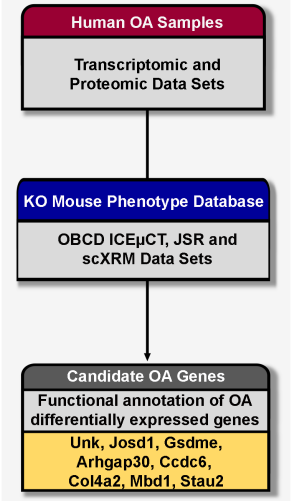

WT
Bhlhe40-/-
Sh3bp4-/-
a
b
c
d
Cg.V (mm3)
0.05
0.04
0.03
0.02
0.01
0.00
*
Bhlhe40-/- Sh3bp4-/-
Median Cg.Th (mm)
0.15
0.10
0.05
0.00
*
**
Bhlhe40-/- Sh3bp4-/-
Cg. Damage Area (%)
10
8
6
4
2
0
*
Bhlhe40-/- Sh3bp4-/-
SC BMC
210
190
170
150
Bhlhe40-/- Sh3bp4-/-

Figure 5

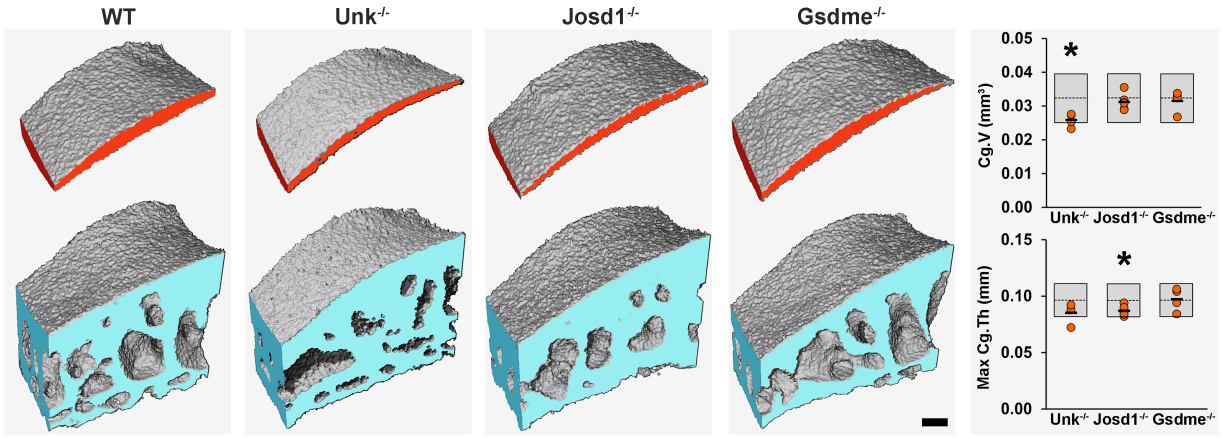
a



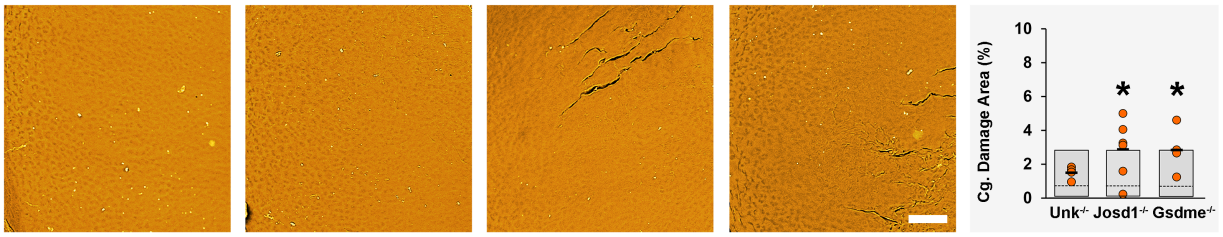
b

Human			
Ensembl ID	Gene	RNA LogFC	Protein LogFC
ENSG00000132478	UNK	-0.17	-0.07
ENSG00000100221	JOSD1	0.08	-
ENSG00000105928	GSDME	0.41	0.12
ENSG00000186517	ARHGAP30	0.69	-
ENSG00000108091	CCDC6	-	0.14
ENSG00000134871	COL4A2	0.45	-0.16
ENSG00000141644	MBD1	-	-0.10
ENSG00000040341	STAU2	0.39	-

c



d



e

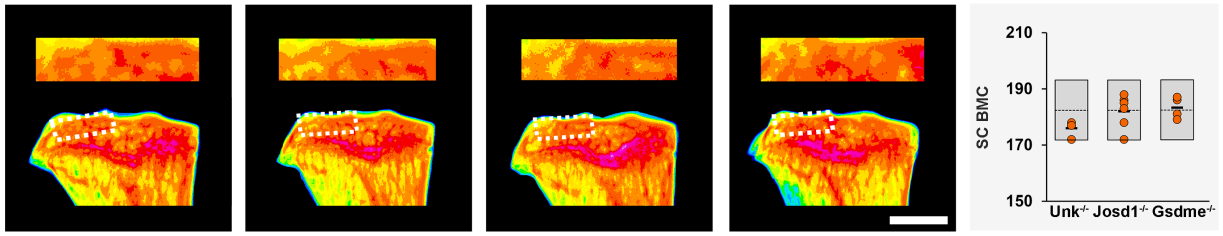


Figure 6

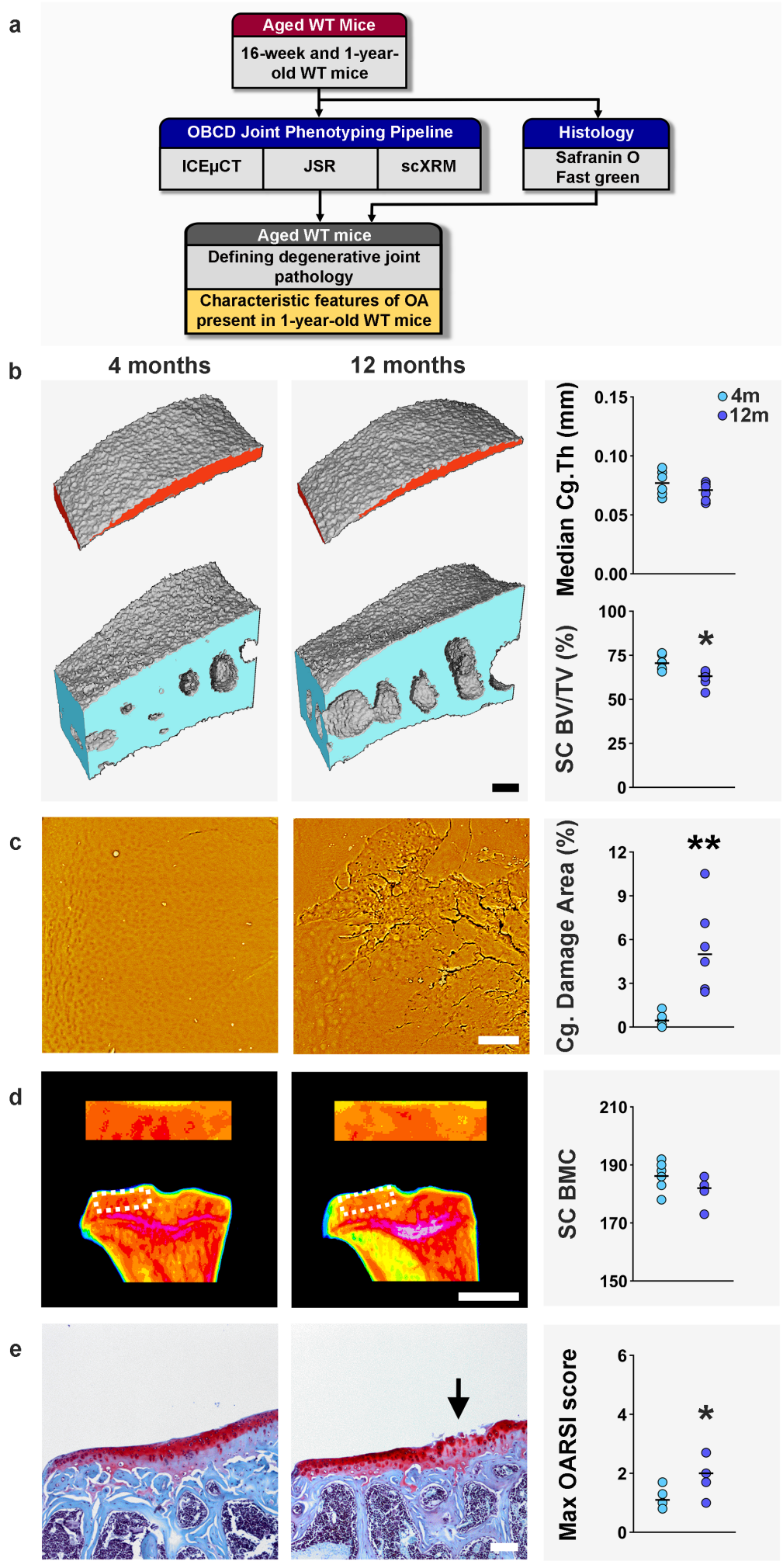
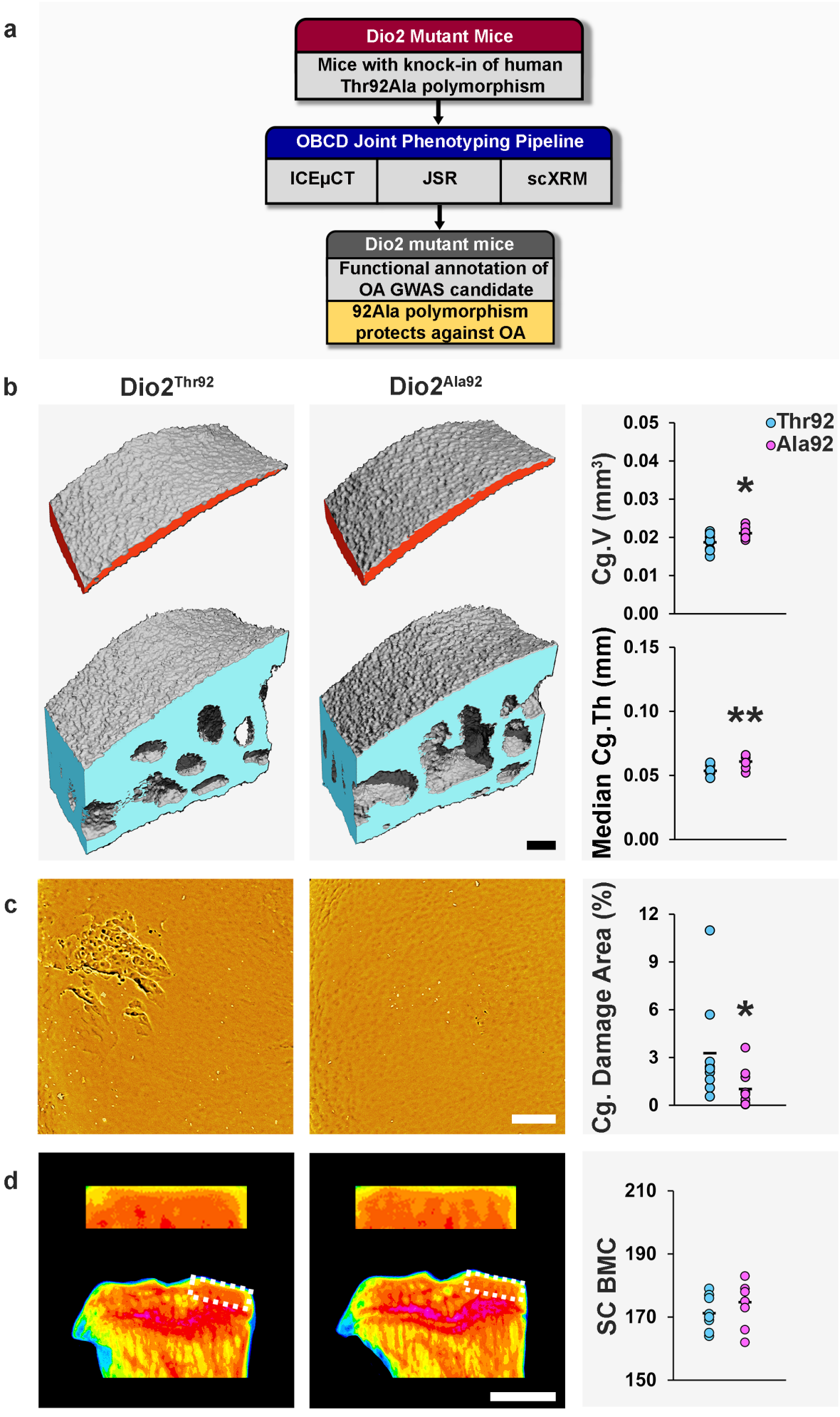

a
Aged WT Mice
16-week and 1-year-old WT mice
OBCD Joint Phenotyping Pipeline
ICEμCT
JSR
scXRM
Histology
Safranin O
Fast green
Aged WT mice
Defining degenerative joint pathology
Characteristic features of OA present in 1-year-old WT mice
b
4 months
12 months
Median Cg.Th (mm)
0.15
0.10
0.05
0.00
4m
12m
SC BV/TV (%)
100
75
50
25
0
*
c
Cg. Damage Area (%)
12
9
6
3
0
**
d
SC BMC
210
190
170
150
e
Max OARSI score
6
4
2
0
*

Figure 7


a
Dio2 Mutant Mice
Mice with knock-in of human Thr92Ala polymorphism
OBCD Joint Phenotyping Pipeline
ICEμCT
JSR
scXRM
Dio2 mutant mice
Functional annotation of OA GWAS candidate
92Ala polymorphism protects against OA
b
Dio2Thr92
Dio2Ala92
Thr92
Ala92
Cg.V (mm3)
0.05
0.04
0.03
0.02
0.01
0.00
*
Median Cg.Th (mm)
0.15
0.10
0.05
0.00
**
c
Cg. Damage Area (%)
12
9
6
3
0
*
d
SC BMC
210
190
170
150

Accelerating functional gene discovery in osteoarthritis

Natalie C. Butterfield[1], Katherine F. Curry[1], Julia Steinberg[2,3,4], Hannah Dewhurst[1], Davide Komla-Ebri[1], Naila S. Mannan[1], Anne-Tounsia Adoum[1], Victoria D. Leitch[1], John G. Logan[1], Julian A. Waung[1], Elena Ghirardello[1], Lorraine Southam[2,3], Scott E. Youlten[5], J, Mark Wilkinson[6,7], Elizabeth A. McAninch[8], Valerie E. Vancollie[3], Fiona Kussy[3], Jacqueline K. White[3,9], Christopher J. Lelliott[3], David J. Adams[3], Richard Jacques[10], Antonio C. Bianco[11], Alan Boyde[12], Eleftheria Zeggini[2,3], Peter I. Croucher[5], Graham R. Williams[1,13*] and J. H. Duncan Bassett[1,13*]

[1]Molecular Endocrinology Laboratory, Department of Metabolism, Digestion and Reproduction, Imperial College London, London W12 0NN, UK.

[2]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany

[3]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

[4]Cancer Council NSW, Sydney, New South Wales 2000, Australia

[5]The Garvan Institute of Medical Research and St. Vincent's Clinical School, University of New South Wales Medicine, Sydney, New South Wales 2010, Australia

[6]Department of Oncology and Metabolism, University of Sheffield, Sheffield S10 2RX, UK

[7]Centre for Integrated Research into Musculoskeletal Ageing and Sheffield Healthy Lifespan Institute, University of Sheffield, Sheffield S10 2TN, UK

[8]Division of Endocrinology and Metabolism, Rush University Medical Center, Chicago, IL 60612, USA

[9]The Jackson Laboratory, Bar Harbor, ME 04609, USA

[10]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield S1 4DA, UK

[11]Section of Adult and Pediatric Endocrinology, Diabetes & Metabolism, Department of Medicine, University of Chicago, Chicago, IL 60637, USA

[12]Dental Physical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

[13]These authors contributed equally

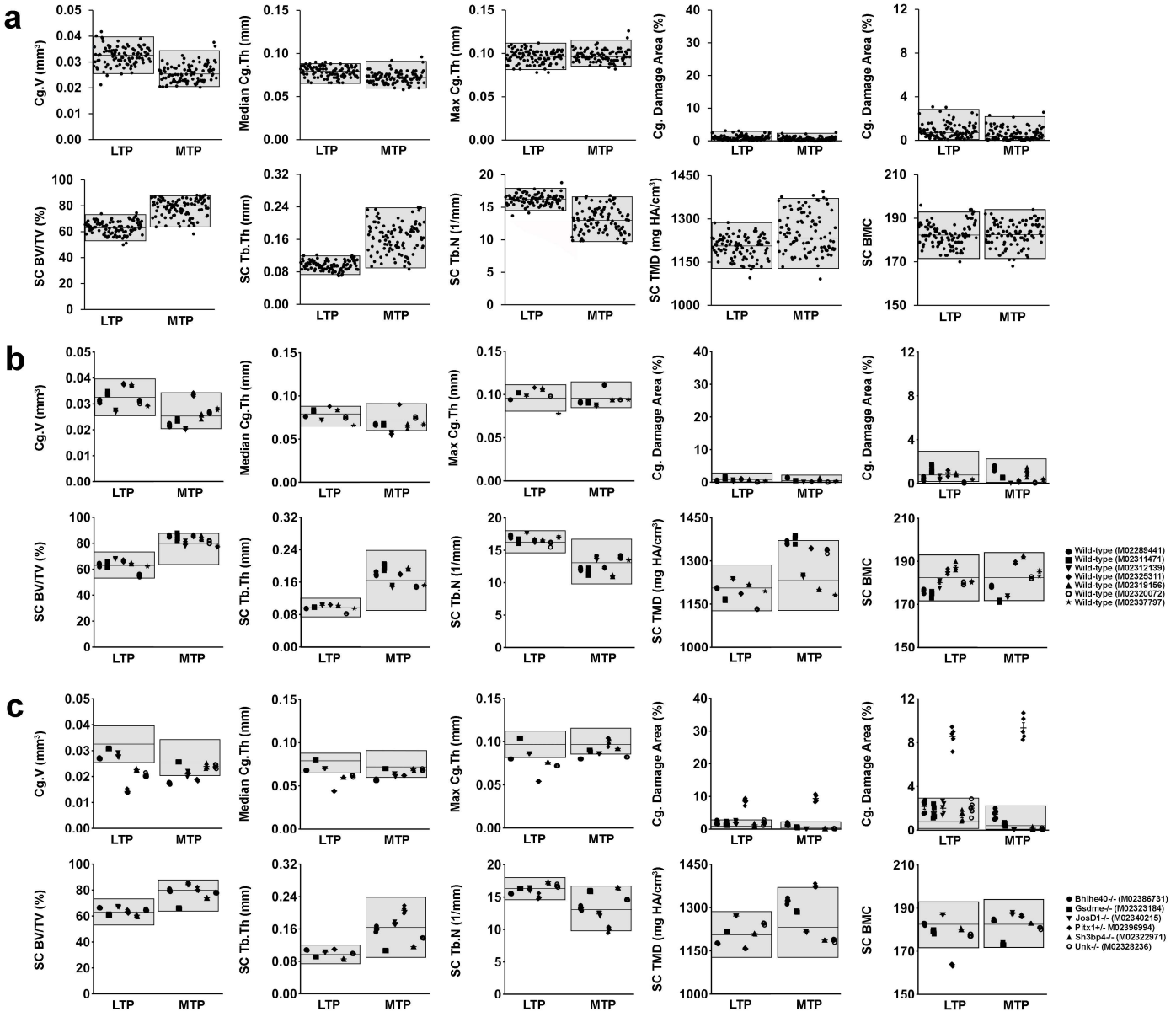
*Corresponding authors

Supplementary Information Contents

Supplementary Figures

Supplementary Methods

Supplementary References



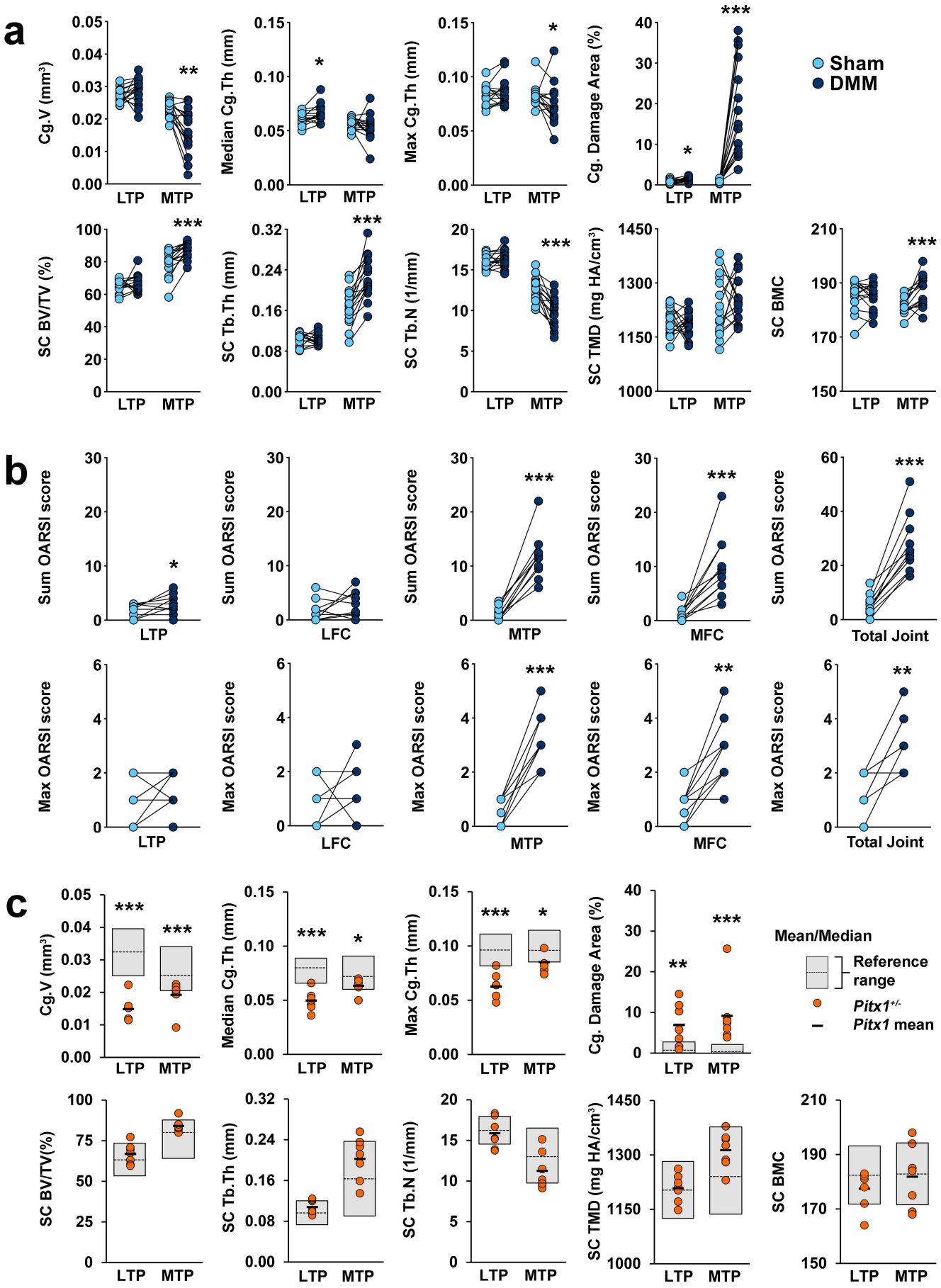
Reference range data for joint phenotype parameters

a. Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum articular cartilage thickness (Max Cg.Th), Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) of 16-week-old male WT mice (n=100). Black dots: individual mice.

b. Repeatability of measurement for each parameter determined by five repeat analyses of seven samples obtained from individual 16-week-old wild-type mice. Black symbols are repeat measures for each sample, with mean ± standard error of the mean.

c. Repeatability of measurement for each parameter determined by five repeat analyses of six samples obtained from individual 16-week-old mutant mice representing the full spectrum of phenotype severity. Black symbols are repeat measures for each sample with mean ± standard error of the mean.

In ***a-c***, grey boxes are reference ranges derived from 100 wild-type samples. For normally distributed parameters [Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP)], reference range is 2 standard deviations above and below the mean (black line). For non-normally distributed parameters [Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP), SC TMD (MTP)], reference range is the 2.5th-97.5th percentile, and black line is the median. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.


a
Cg.V (mm³)
Median Cg.Th (mm)
Max Cg.Th (mm)
Cg. Damage Area (%)
Sham
DMM
SC BV/TV (%)
SC Tb.Th (mm)
SC Tb.N (1/mm)
SC TMD (mg HA/cm³)
SC BMC
LTP
MTP
b
Sum OARSI score
Max OARSI score
LTP
LFC
MTP
MFC
Total Joint
c
Cg.V (mm³)
Median Cg.Th (mm)
Max Cg.Th (mm)
Cg. Damage Area (%)
Mean/Median
Reference range
Pitx1+/-
Pitx1 mean
SC BV/TV(%)
SC Tb.Th (mm)
SC Tb.N (1/mm)
SC TMD (mg HA/cm³)
SC BMC
LTP
MTP

Supplementary Figure 2

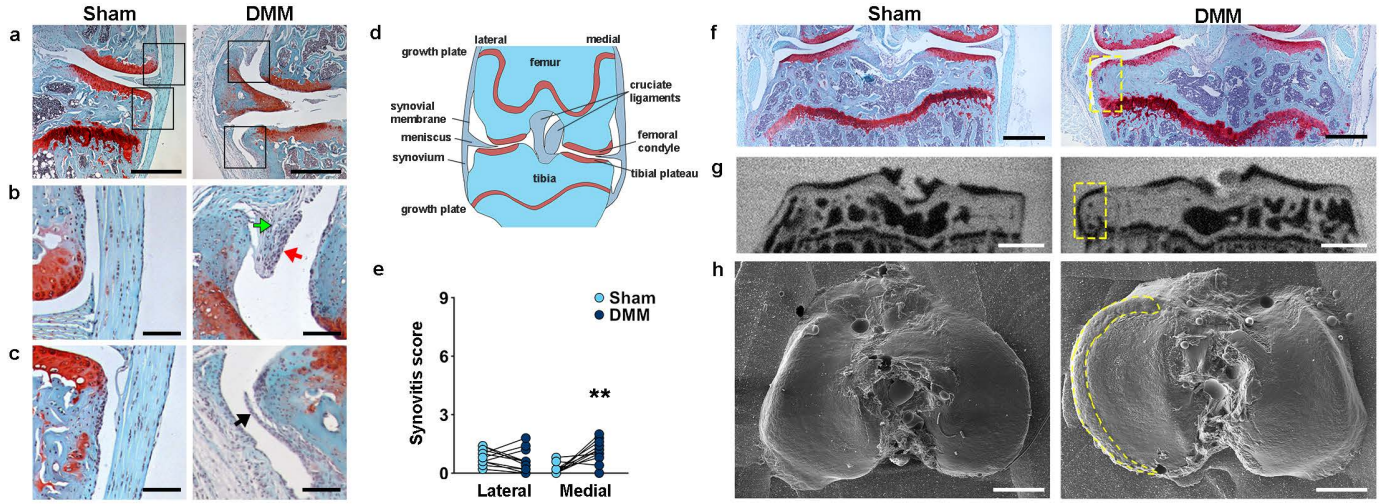
Joint abnormalities following destabilization of the medial meniscus (DMM) surgery and severe early onset osteoarthritis in Pitx1$^{+/-}$ mice

a. Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum articular cartilage thickness (Max Cg.Th), Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) of sham and DMM-operated knees of wild type (WT) mice 12 weeks after surgery (n=16). Paired data (sham versus DMM) are shown for each mouse. **P*<0.05, ***P*<0.01, 2-tailed Wilcoxon matched pairs signed rank test for Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP) and SC TMD (MTP), or 2-tailed paired *t*-test for Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP). Cg.V (MTP): *P*=0.001312, Median Cg.Th (LTP): *P*=0.021729, Max. Cg.Th (MTP): *P*=0.017334, Cg. Damage Area (LTP): *P*=0.013092, (MTP): *P*=0.000031, SC BV/TV (MTP): *P*=0.000061, SC Tb.Th (MTP): *P*=0.000013, SC Tb.N (MTP): *P*=0.000031, SC BMC (MTP): *P*=0.000477.

b. Graphs showing Osteoarthritis Research Society International (OARSI) histological scores in sham and DMM-operated knees of WT mice 12 weeks after surgery (n=11). Parameters include the summed and maximum scores on the LTP, lateral femoral condyle (LFC), MTP, medial femoral condyle (MFC), and combined values for the total joint. **P*<0.05, ***P*<0.01, 2-tailed Wilcoxon matched pairs signed rank test. Sum OARSI scores: LTP: *P*=0.0469, MTP: *P*=0.000977, MFC: *P*=0.000977, total joint: *P*=0.000977. Max. OARSI scores: MTP: *P*=0.000977, MFC: *P*=0.001953, total joint: *P*=0.001953.

c. Graphs showing articular cartilage and subchondral bone parameters in the LTP and MTP of 16-week-old *Pitx1*$^{+/-}$ mice (n=7). Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters [Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP)], reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters [Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP) SC TMD (MTP)], reference range is the 2.5th-97.5th percentile, and dashed line is the median. **P*<0.00568, ***P*<0.001, ****P*<0.0001, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. Cg.V (LTP): *P*=0.00001, (MTP): *P*=0.00008, Median Cg.Th (LTP): *P*=0.00001, (MTP): *P*=0.00307 Max. Cg.Th (LTP): *P*=0.00001, (MTP): *P*=0.00319, Cg. Damage Area (LTP): *P*=0.00014 (MTP): *P*=0.00001. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.

Supplementary Figure 3



Synovitis and osteophyte formation in destabilization of the medial meniscus (DMM)-operated mice

a. Example Safranin O/Fast green-stained coronal sections from sham and DMM-operated joints demonstrating the signs of synovitis in the operated knee. Black boxes indicate regions of interest shown in ***b*** and ***c***.

b. Synovial hyperplasia (red arrow, 10/11 knees) and sub-synovial inflammation with new blood vessel formation (green arrow, 8/11 knees) are shown in the DMM-operated knee.

c. Pannus formation (black arrow, 1/11 knee) in the DMM-operated knee.

d. Schematic representation of the mouse knee joint, red; cartilage, blue; bone, grey; soft tissue

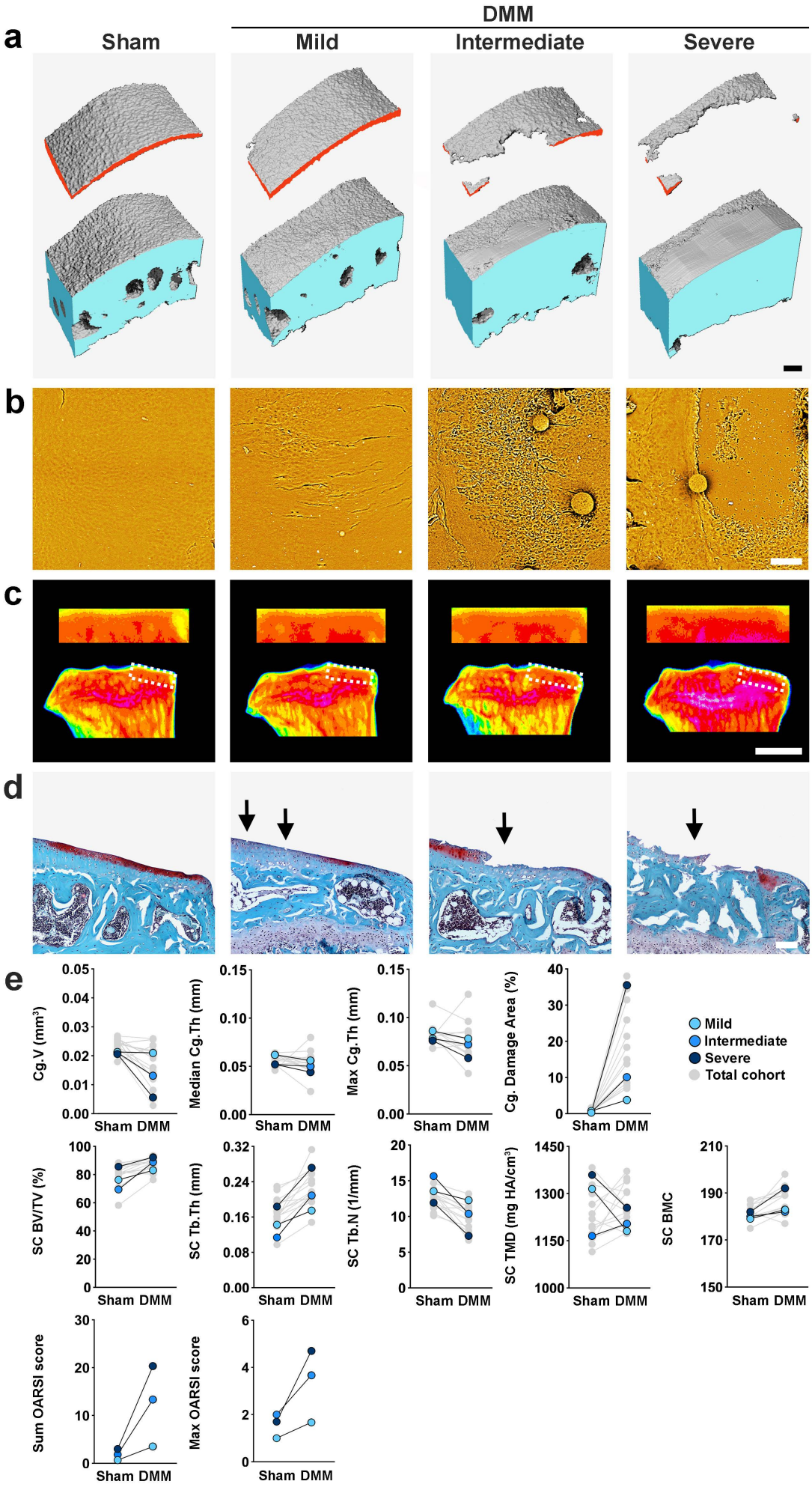
e. Mean synovitis scores in the lateral and medial compartments of sham and DMM-operated knees (n=11, ***P*= 0.0059, 2-tailed Wilcoxon matched-pairs signed rank test). Source data are provided as a Source Data file.

f. Safranin O/Fast green-stained coronal sections from sham and DMM-operated joints showing osteophyte formation (yellow box).

g. Iodine contrast-enhanced µCT (ICEµCT) images showing osteophyte formation (yellow box).

h. Joint surface replica (JSR) images showing osteophyte formation outlined in yellow. Scale bars = 500µm (panels a, f, g, h), and 100µm (panels b, c).

Supplementary Figure 4



Supplementary Figure 4

Validation of imaging methods

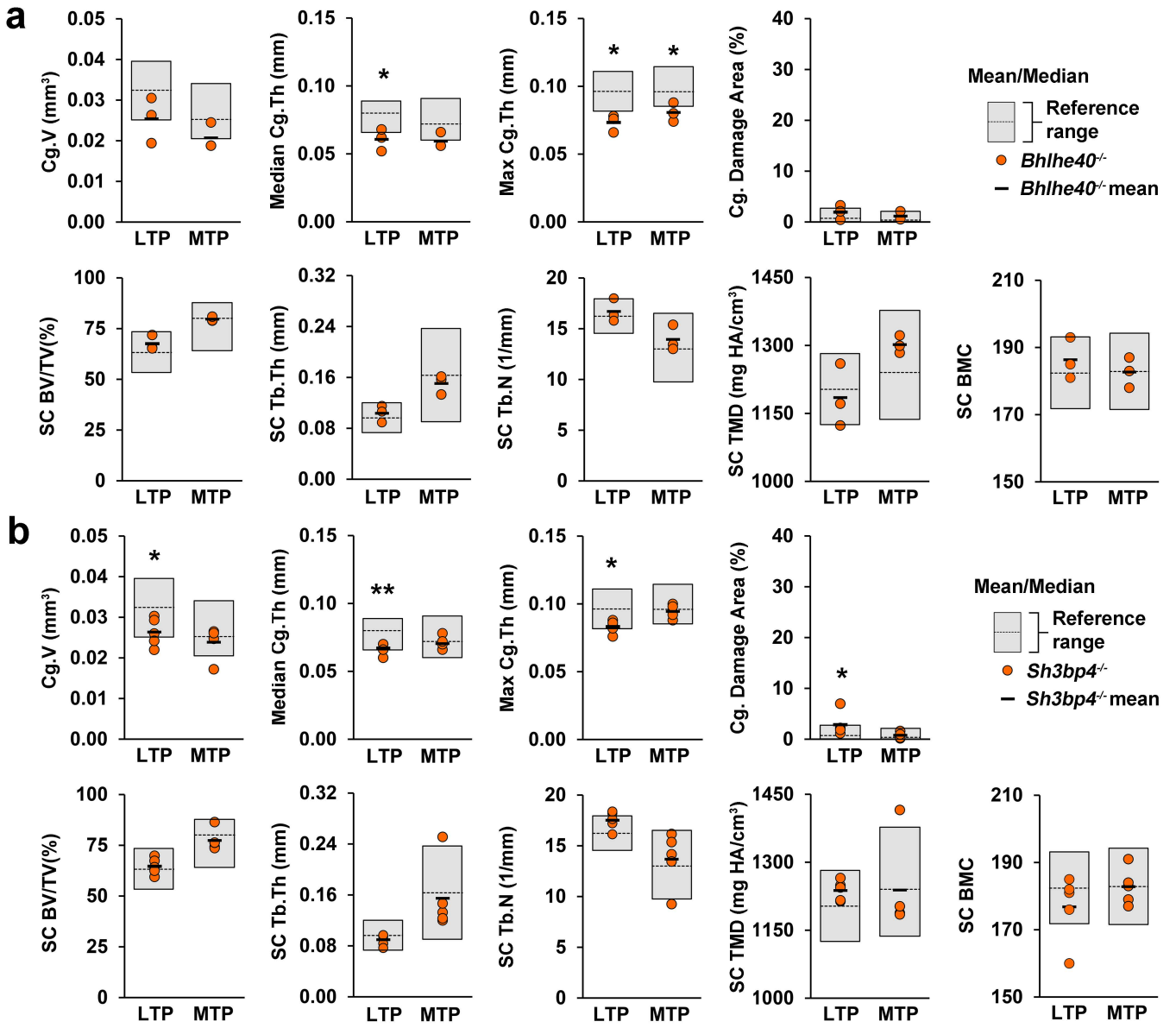
a. Iodine contrast enhanced μCT images of medial tibial plateau articular cartilage (red) and subchondral bone (blue) from 22-week-old wild type (WT) male mice 12 weeks after sham operation or following destabilization of the medial meniscus (DMM) surgery that resulted in either mild, intermediate or severe osteoarthritis (OA).

b. Back-scattered electron scanning-electron microscopy images of medial tibial plateau joint surface replicas from sham and DMM-operated WT mice with mild, intermediate or severe OA.

c. X-ray microradiography images of proximal tibia and the medial tibial plateau subchondral bone region of interest (dashed box) from sham and DMM-operated WT mice with mild, intermediate or severe OA. Greyscale images are pseudocoloured according to a 16 colour look-up table in which low bone mineral content (BMC) is yellow and high BMC is pink.

d. Coronal sections of medial tibial plateaux stained with Safranin-O/Fast green from three sham and three DMM-operated WT mice with mild, intermediate or severe OA (Sections assessed at 5 levels through each knee joint). Arrows indicate areas of cartilage damage. Scale bars = 100μm (a, b, d) and 1mm (c).

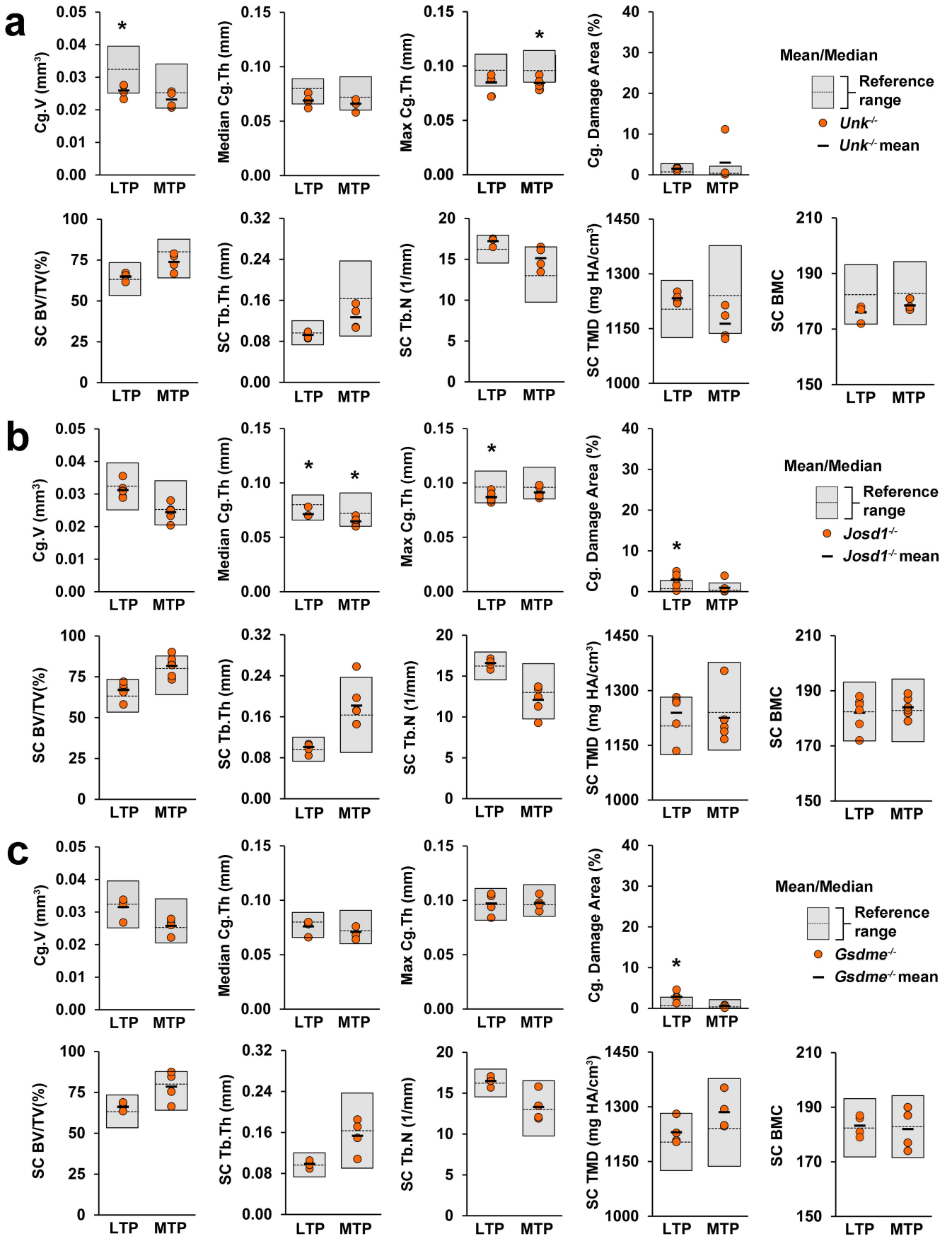
e. Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum articular cartilage thickness (Max Cg.Th) and Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters together with summed and maximum Osteoarthritis Research Society International (OARSI) histological scores on the medial tibial plateaux of sham and DMM-operated knees from three mice with mild, intermediate and severe OA. Results from the entire cohort in grey, with mild, intermediate and severe examples coloured according to key. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.



Early onset osteoarthritis in Bhlhe40$^{-/-}$ and Sh3pb4$^{-/-}$ mutant mice

a. Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum articular cartilage thickness (Max Cg.Th) and Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) from 16-week-old *Bhlhe40$^{-/-}$* mice (n=3).

b. Graphs showing articular cartilage and subchondral bone parameters in the LTP and MTP from 16-week-old *Sh3pb4$^{-/-}$* mice (n=5). Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters [Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP)], reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters [Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP), SC TMD (MTP)], reference range is the 2.5th-97.5th percentile, and dashed line is the median. *P<0.00568, **P<0.001, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. *Bhlhe40$^{-/-}$*: Median Cg.Th (LTP): P=0.00469, Max. Cg.Th (LTP): P=0.00347, Max. Cg.Th (MTP): P=0.00498. *Sh3bp4$^{-/-}$*: Cg.V (LTP): P=0.00311, Median Cg.Th (LTP): P=0.00081. Max. Cg.Th (LTP): P=0.00133, Cg. Damage Area (LTP): P=0.00282. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.


a
b
c
Cg.V (mm³)
Median Cg.Th (mm)
Max Cg.Th (mm)
Cg. Damage Area (%)
SC BV/TV(%)
SC Tb.Th (mm)
SC Tb.N (1/mm)
SC TMD (mg HA/cm³)
SC BMC
LTP
MTP
Mean/Median
Reference range
Unk-/-
Unk-/- mean
Josd1-/-
Josd1-/- mean
Gsdme-/-
Gsdme-/- mean

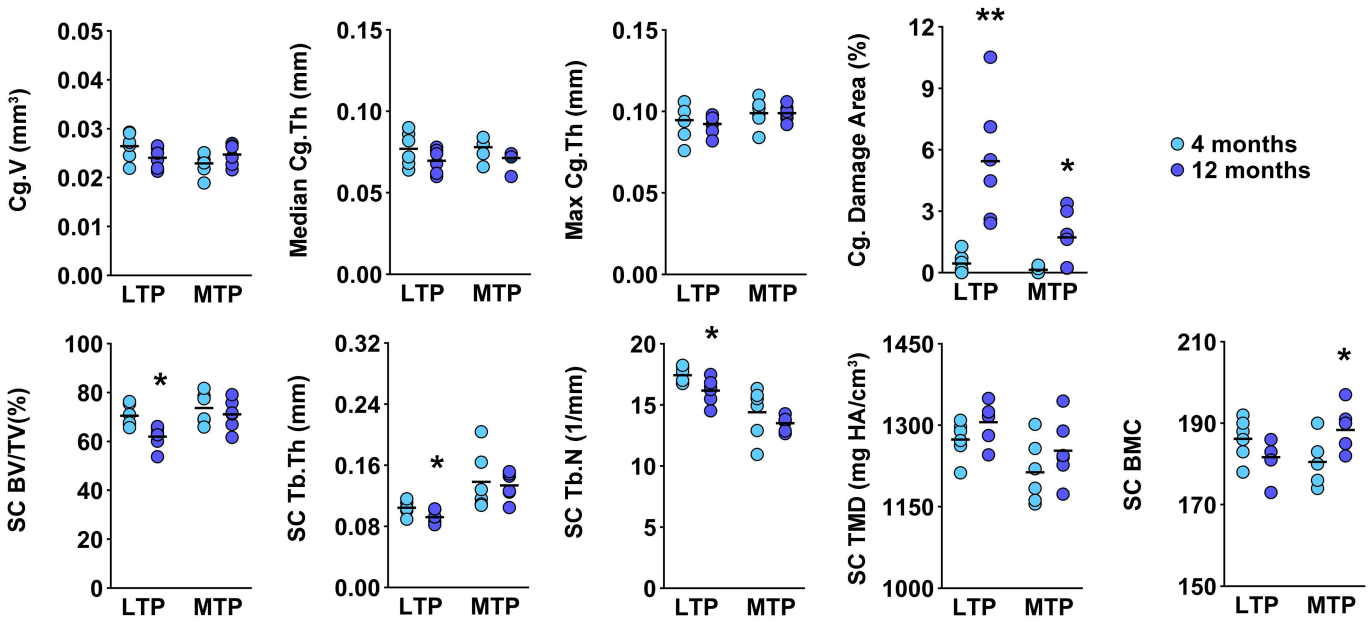
Early onset osteoarthritis in mice with deletion of genes differentially expressed in human osteoarthritis cartilage

a. Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum Cg.Th (Max Cg.Th), Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) of 16-week-old *Unk*$^{-/-}$ mice (n=4).

b. Graphs showing articular cartilage and subchondral bone parameters in the LTP and MTP of 16-week-old *Josd1*$^{-/-}$ mice (n=6).

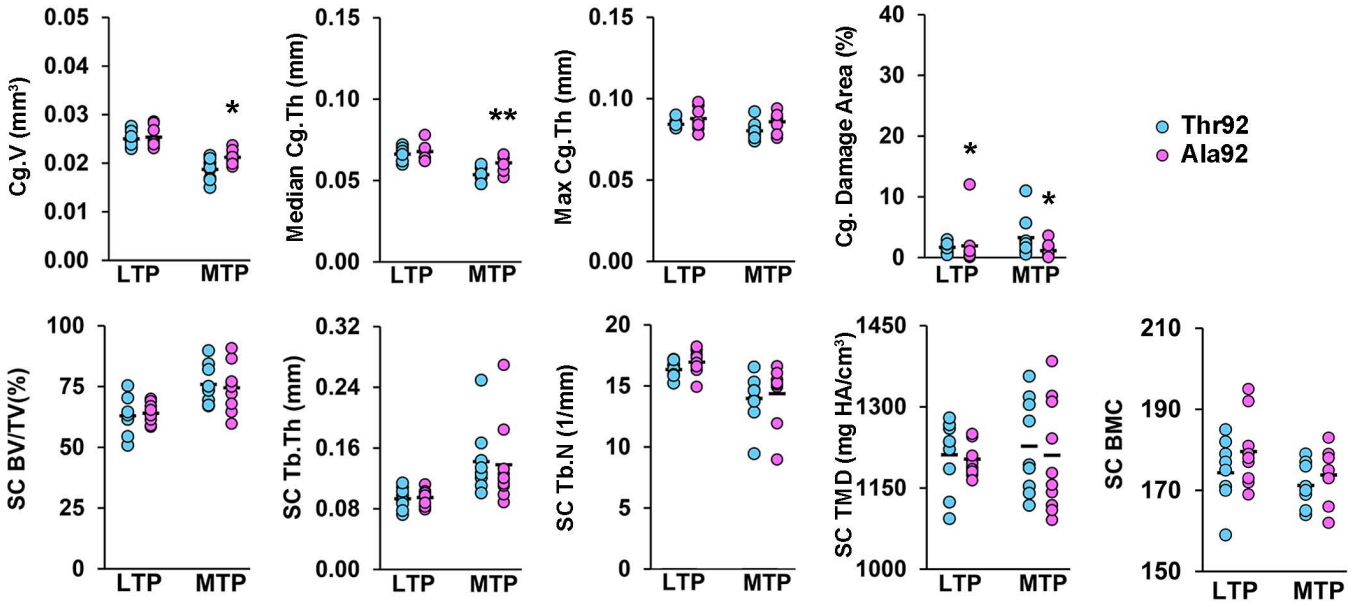
c. Graphs showing articular cartilage and subchondral bone parameters in the LTP and MTP of 16-week-old *Gsdme*$^{-/-}$ mice (n=4). Orange circles: individual mutant samples, black horizontal lines: sample mean. Grey boxes: reference ranges derived from 100 wild-type samples. For normally distributed parameters [Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP)], reference range is 2 standard deviations above and below the mean (dashed line). For non-normally distributed parameters [Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP), SC TMD (MTP)], reference range is the 2.5th-97.5th percentile, and dashed line is the median. **P*<0.00568, Bonferroni-corrected 2-tailed Wilcoxon rank sum test. *Unk*$^{-/-}$*:* Cg.V (LTP): *P*=0.00215, Max Cg.Th (MTP): *P*=0.00295. *Josd*$^{-/-}$*:* Median Cg.Th (LTP): *P*=0.00344, (MTP): *P*=0.00487, Max Cg.Th (LTP): *P*=0.00362, Cg. Damage area (LTP): *P*=0.00563. *Gsdme*$^{-/-}$: Cg. Damage Area (LTP): *P*=0.00355. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.

Supplementary Figure 7



Age-related joint degeneration

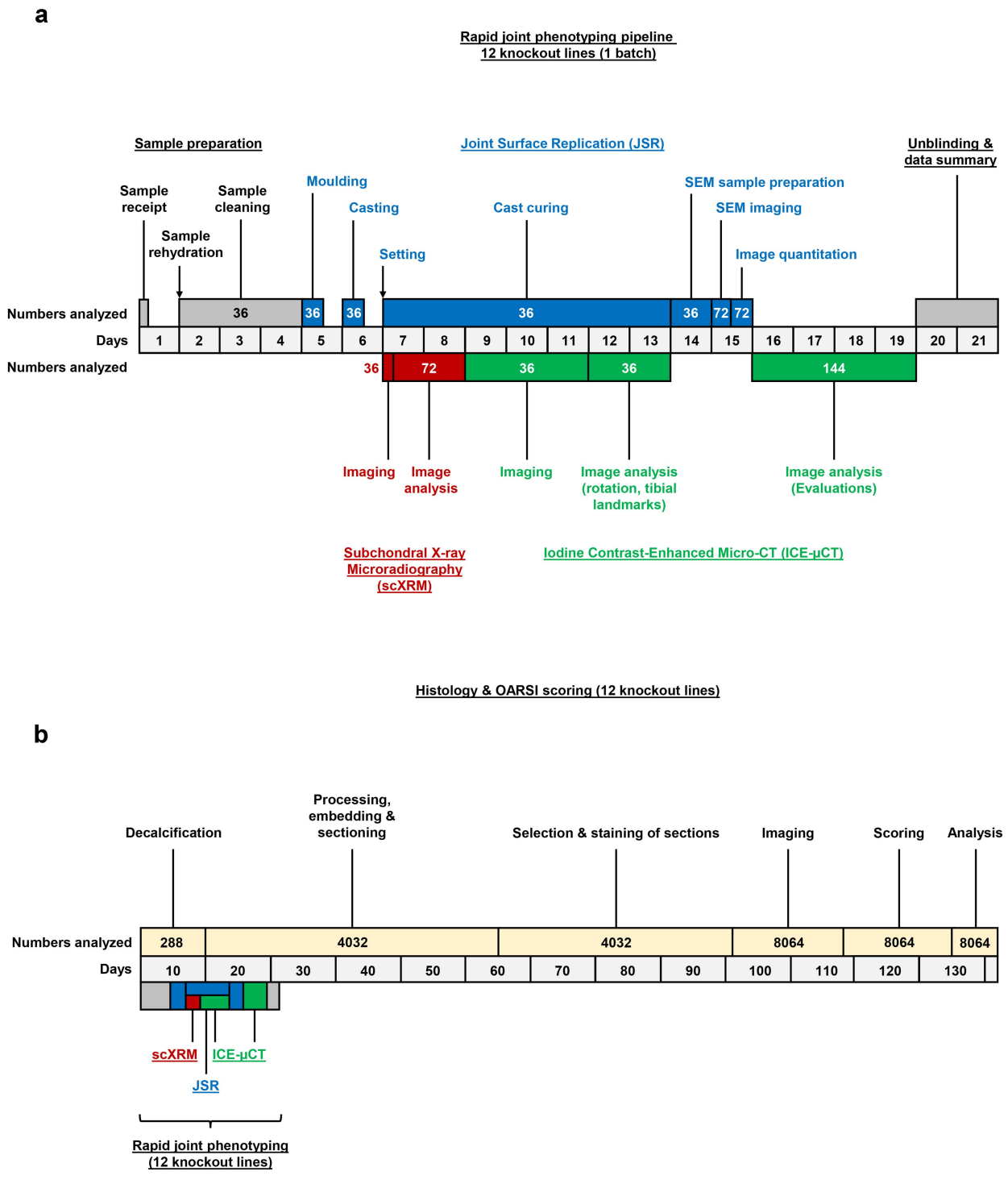
Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum Cg.Th (Max Cg.Th), Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) of 4-month-old (n=6) and 12-month-old (n=6) WT mice. **P*<0.05, ***P*<0.01, 2-tailed Wilcoxon rank sum test for Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP) and SC TMD (MTP), or 2-tailed *t*-test for Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP) and SC BMC (LTP, MTP). Cg. Damage Area (LTP): *P*=0.002, (MTP): *P*=0.026, SC BV/TV (LTP): *P*=0.008, SC Tb.Th (LTP): *P*=0.039, SC Tb.N (LTP): *P*=0.035, SC BMC (MTP): *P*=0.034. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.



Mice with a Dio2^{Ala92} polymorphism are protected from osteoarthritis

Graphs showing articular cartilage (articular cartilage volume (Cg.V), median articular cartilage thickness (Median Cg.Th), maximum articular cartilage thickness (Max Cg.Th), Cg. damage area) and subchondral bone (subchondral bone volume per tissue volume (SC BV/TV), SC trabecular thickness (SC Tb.Th), SC trabecular number (SC Tb.N), SC tissue mineral density (SC TMD), SC bone mineral content (SC BMC)) parameters in the lateral and medial tibial plateaux (LTP, MTP) from 16-week-old *Dio2*^{Thr92} (n=9) and *Dio2*^{Ala92} (n=10) mice. **P*<0.05, ***P*<0.01, 2-tailed Wilcoxon rank sum test for Cg.V (MTP), Median Cg.Th (LTP, MTP), Max Cg.Th (MTP), Cg. Damage Area (LTP, MTP), SC-BV/TV (MTP), SC Tb.N (MTP) and SC TMD (MTP), or 2-tailed *t*-test for Cg.V (LTP), Max Cg.Th (LTP), SC-BV/TV (LTP), SC Tb.Th (LTP, MTP), SC Tb.N (LTP), SC TMD (LTP), SC BMC (LTP, MTP). Cg.V (MTP): *P*=0.019, Median Cg.Th (MTP): *P*=0.006, Cg. Damage Area (LTP): *P*=0.043, (MTP): *P*=0.016. mm; millimeters, mg HA/cm^3; milligrams of hydroxyapatite/cubic centimeter. Source data are provided as a Source Data file.

Supplementary Figure 9

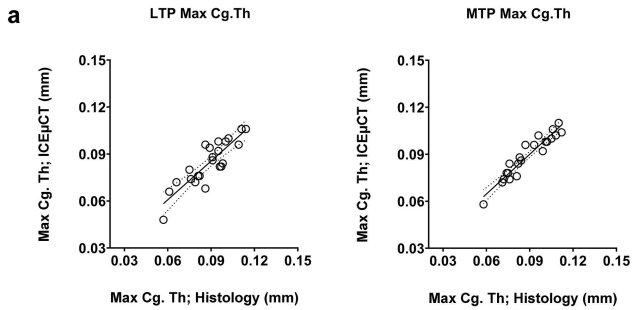


Origins of Bone and Cartilage Disease (OBCD) rapid-throughput joint phenotyping compared to Osteoarthritis Research Society International (OARSI) histological scoring

a. Timeline showing number of days and number of samples and images analysed for OBCD rapid-throughput joint phenotyping of 12 mutant mouse lines (n=3 mice required per line).

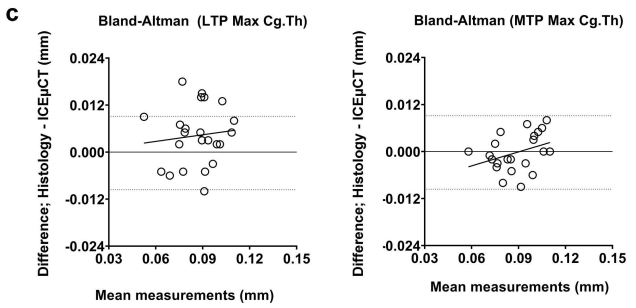
b. Timeline showing number of days and number of samples and histological sections analysed for phenotyping 12 mutant mouse lines by gold-standard OARSI scoring (n=12 mutant and 12 wild type mice required per line).

Supplementary Figure 10



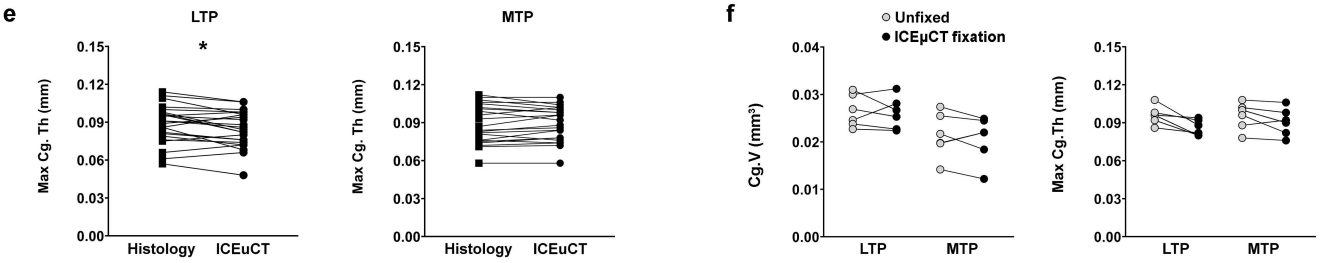
b

Linear regression	Medial Tibial Plateau (MTP)	Lateral Tibial Plateau (LTP)
Slope	0.847	0.824
R^2	0.905	0.755
P-value	<0.0001	<0.0001
Deviation from zero?	Significant (***)	Significant (***)
Equation	Y = 0.8469*X + 0.01380	Y = 0.8244*X + 0.01133



d

Bland-Altman plots		Medial Tibial Plateau (MTP)	Lateral Tibial Plateau (LTP)
Bias		-0.0002273 mm	0.004217 mm
SD of bias		0.00478 mm	0.007556 mm
95% limits of agreement	From	-0.009597 mm	-0.01059 mm
	To	0.009142 mm	0.01903 mm
Trendline			
R^2		0.125	0.011
Slope		0.119	0.056
P-value		0.106	0.630
Deviation from zero?		Not Significant	Not Significant
Equation		Y = 0.1190*X - 0.01080	Y = 0.05639*X - 0.0006577



Validation of iodine contrast-enhanced μCT (ICEμCT)

a. Maximum cartilage thickness (max. Cg.Th, mm) determined by ICEμCT versus maximum cartilage thickness determined by histology (mm). Lateral tibial plateau (LTP); medial tibial plateau (MTP); linear regression and 95% confidence intervals.

b. Correlation (R^2) between cartilage thickness determined by ICEμCT and histology.

c. Bland-Altman plots comparing maximum cartilage thickness determined by ICEμCT versus histology. Linear regression and 95% limits of agreement are shown.

d. Table demonstrating that Bland-Altman analysis does not deviate from zero.

e. Max Cg.Th is significantly larger when measured by histology than when measured by ICEμCT, for LTP only (P=0.014, paired 2-tailed *t*-test assuming unequal variance, **P*<0.05).

f. Articular cartilage volume (Cg.V) and Max Cg.Th in samples under ICEμCT conditions (fixation in 10% neutral buffered formalin, stored in 70% ethanol, rehydrated 24-36h in phosphate buffered saline) are not significantly different from that of unfixed cartilage (Bonferroni-corrected *P*-value **P*<0.017). Cg.V (LTP): *P*=0.705, (MTP): *P*=0.500, Max. Cg.Th (LTP): *P*=0.0189, (MTP): *P*=0.135. mm; millimeter. Source data are provided as a Source Data file.

Supplementary Methods

Isolation of chondrocytes

For cohorts 1-2, the isolation of chondrocytes has been described previously[1] and the same protocol was followed for cohort 4. The protocol is described below. Osteochondral samples were transported in Dulbecco's modified Eagle's medium (DMEM)/F-12 (1:1) (Life Technologies) supplemented with 2mM glutamine (Life Technologies), 100U/mL penicillin, 100µg/mL streptomycin (Life Technologies), 2.5µg/mL amphotericin B (Sigma-Aldrich) and 50µg/mL ascorbic acid (Sigma-Aldrich) (serum-free media). Half of each sample was then taken forward for chondrocyte extraction. Cartilage was removed from the bone, dissected and washed twice in 1xPBS. Tissue was digested in 3mg/mL collagenase type I (Sigma-Aldrich) in serum-free media overnight at 37°C on a flatbed shaker. The resulting cell suspension was passed through a 70µm cell strainer (Fisher Scientific) and centrifuged at 400g for 10 minutes. Subsequently, the cell pellet was washed twice in serum-free media and centrifuged at 400g for 10 minutes. The resulting cell pellet was resuspended in serum-free media. Cells were counted using a haemocytometer and the viability checked using trypan blue exclusion (Invitrogen). The optimal cell number for spin column extraction from cells was between $4x10^6$ and $1x10^7$. Cells were then pelleted and homogenized.

For cohort 3, the extraction of chondrocytes has previously been described[2]. The protocol was based on that for cohorts 1, 2 and 4 above, and is highly similar, with differences summarized in the following. Each cartilage portion was minced with a scalpel and placed in 20mL of Dulbecco’s modified Eagle medium (Invitrogen) containing 10% foetal bovine serum (Invitrogen) and 6mg/mL collagenase A (Sigma).

The tissue culture flasks were incubated overnight to digest the cartilage pieces. The resulting cell suspension was passed through a 30µm filter (Miltenyi) and centrifuged at 400g for 10 minutes. The cell pellet was then re-suspended in 1mL of PBS and counted on a haemocytometer following 1:1 mixing with trypan blue to determine cell viability.

DNA, RNA and protein extraction

DNA, RNA, and protein extraction was carried out using Qiagen AllPrep DNA/RNA/Protein Mini Kit following manufacturer's instructions. For cohort 3, minor modifications were made to the approach as previously described[2] and the protocol is summarized below. Cells were lysed in 350µL RLT buffer containing 3.5µL β-mercaptoethanol and centrifuged in a Qiashredder column (Qiagen) for 2min at 20,800g. RNA, DNA and protein were purified using a series of column based centrifugation steps and RNA was eluted in 34µL RNase-free water, DNA in 100µL EB buffer and the protein pellet solubilized in 100µL 10% sodium dodecyl sulphate in PBS. DNA and RNA preparations were re-precipitated to remove buffer salts, and to concentrate the final product 20µL 3M sodium acetate, pH 5.5 (Ambion) was added per 100µL of DNA or RNA solution and mixed gently. Then, 283µL of molecular grade ethanol (Sigma) was added per 100µL of DNA or RNA solution to precipitate the nucleic acids and the mixture placed at -20°C overnight. Samples were centrifuged at 20,800g for 30 minutes and the supernatant removed, the pellet was washed with 900µL 70% ethanol and the nucleic acids re-pelleted by spinning at 20,800g for 10 minutes. Nucleic acids were resuspended in 32µL DNase/RNase free water. Samples were frozen at -80°C (cohorts 1, 2, 4) or -70°C (cohort 3) prior to assays.

RNA sequencing

We purified poly-A tailed RNA (mRNA) from total RNA using Illumina's TruSeq RNA Sample Prep v2 kits. We then fragmented the mRNA using metal ion-catalyzed hydrolysis and synthesized a random-primed cDNA library. The resulting double-strand cDNA was used as the input to a standard Illumina library prep, whereby ends were repaired to produce blunt ends by a combination of fill-in reactions and exonuclease activity. We performed A-tailing to allow samples to be pooled, by adding an “A” base to the blunt ends and ligation to Illumina Paired-end Sequencing adapters containing unique index sequences. Due to better performance, the 10-cycle PCR amplification of libraries was carried out using KAPA Hifi Polymerase. A post-PCR Agilent Bioanalyzer was used to quantify samples, followed by sample pooling and size-selection of pools using the LabChip XT Caliper. The multiplexed libraries were sequenced on the Illumina HiSeq 2000 for cohort 1 and HiSeq 4000 for cohorts 2-4 (75bp paired-ends). Sequenced data underwent initial analysis and quality control on reads as standard. This yielded gene expression data for samples from 110 patients.

Proteomics

Proteomics analysis was performed on cartilage samples from 103 patients. For cohort 1, all steps of protein digestion, 6-plex TMT labelling, peptide fractionation and LC-MS analysis on the Dionex Ultimate 3000 UHPLC system coupled with the high-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Scientific), were previously described[1] and details are given below. The sample preparation protocol formed the basis of processing for cohorts 2-4 using 10-plex TMT labelling and an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific) with otherwise only minor alterations, as described below.

Proteomics: Cohort 1

Protein digestion and TMT labeling

The protein content of each sample was precipitated by the addition of 30μL TCA 8M at 4°C for 30 min. The protein pellets were washed twice with ice cold acetone and finally re-suspended in 40μL 0.1M triethylammonium bicarbonate, 0.05% SDS with pulsed probe sonication. Protein concentration was measured with Quick Start Bradford Protein Assay (Bio-Rad) according to manufacturer's instructions. Aliquots containing 30μg of total protein were prepared for trypsin digestion. Cysteine disulfide bonds were reduced by the addition of 2μL 50mM tris-2-carboxymethyl phosphine (TCEP) followed by 1h incubation in a heating block at 60°C. Cysteine residues were blocked by the addition of 1μL 200mM freshly prepared Iodoacetamide (IAA) solution and 30min incubation at room temperature in the dark. Trypsin (Pierce, MS grade) solution was added to a final concentration 70ng/μL to each sample for overnight digestion. After proteolysis the peptide samples were diluted to 100μL with 0.1M TEAB buffer. A 41μL volume of anhydrous acetonitrile was added to each TMT 6-plex reagent (Thermo Scientific) vial and after vortex mixing the content of each TMT vial was transferred to each sample tube. Labeling reaction was quenched with 8μL 5% hydroxylamine for 15 min after 1h incubation at room temperature. Samples were pooled and the mixture was dried with SpeedVac concentrator and stored at -20°C until high-pH Reverse Phase (RP) fractionation.

Peptide fractionation

Offline peptide fractionation based on high pH Reverse Phase (RP) chromatography was performed using the Waters, XBridge C18 column (2.1 x 150mm, 3.5μm, 120Å) on a Dionex Ultimate 3000 HPLC system equipped with autosampler. Mobile phase (A) was composed of 0.1% ammonium hydroxide and mobile phase (B) was

composed of 100% acetonitrile, 0.1% ammonium hydroxide. The TMT labelled peptide mixture was reconstituted in 100μL mobile phase (A), centrifuged and injected for fractionation. The multi-step gradient elution method at 0.2mL/min was as follows: for 5 minutes isocratic at 5% (B), for 35min gradient to 35% (B), gradient to 80% (B) in 5min, isocratic for 5minutes and re-equilibration to 5% (B). Signal was recorded at 280nm and fractions were collected in a time-dependent manner every one minute. The collected fractions were dried with SpeedVac concentrator and stored at -20°C until LC-MS analysis.

LC-MS analysis

LC-MS analysis was performed on the Dionex Ultimate 3000 UHPLC system coupled with the high-resolution LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). Each peptide fraction was reconstituted in 40μL 0.1% formic acid and a volume of 10μL was loaded to the Acclaim PepMap 100, 100μm × 2cm C18, 5μm, 100Å trapping column with a user modified injection method at 10μL/min flow rate. The sample was then subjected to a multi-step gradient elution on the Acclaim PepMap RSLC (75μm × 50cm, 2μm, 100Å) C18 capillary column (Dionex) retrofitted to an electrospray emitter (New Objective, FS360-20-10-N-20-C12) at 45°C. Mobile phase (A) was composed of 96% H_2O, 4% DMSO, 0.1% formic acid and mobile phase (B) was composed of 80% acetonitrile, 16% H_2O, 4% DMSO, 0.1% formic acid. The gradient separation method at flow rate 300nL/min was as follows: for 95min gradient to 45% B, for 5min up to 95% B, for 8min isocratic at 95% B, re-equilibration to 5% B in 2min, for 10min isocratic at 5% B.

The ten most abundant multiply charged precursors within 380-1500m/z were selected with FT mass resolution of 30,000 and isolated for HCD fragmentation with isolation

width 1.2Th. Normalized collision energy was set at 40 and the activation time was 0.1ms for one microscan. Tandem mass spectra were acquired with FT resolution of 7,500 and targeted precursors were dynamically excluded for further isolation and activation for 40 seconds with 10ppm mass tolerance. FT max ion time for full MS experiments was set at 200ms and FT MSn max ion time was set at 100ms. The AGC target vales were 3×10^6 for full FTMS and 1×10^5 for MSn FTMS. The DMSO signal at m/z 401.922718 was used as a lock mass.

Proteomics: Cohorts 2-4

Protein digestion and TMT labeling

The protein content of each sample was precipitated by the addition of 30µL TCA 8M at 4°C for 30min. The protein pellets were washed twice with ice cold acetone and finally re-suspended in 40µL 0.1M triethylammonium bicarbonate, 0.1% SDS with pulsed probe sonication. Equal aliquots containing at least 10µg of total protein were reduced with 5mM TCEP for 1h at 60°C, alkylated with 10mM Iodoacetamide and subjected to overnight trypsin (70ng/µL) digestion. TMT 10-plex (Thermo Scientific) labelling was performed according to manufacturer's instructions at equal amounts of tryptic digests. Samples were pooled and the mixture was dried with SpeedVac concentrator and stored at -20°C until peptide fractionation.

Peptide fractionation

Offline peptide fractionation was based on high pH Reverse Phase (RP) chromatography using the Waters, XBridge C18 column (2.1 x 150mm, 3.5µm) on a Dionex Ultimate 3000 HPLC system. Mobile phase A was 0.1% ammonium hydroxide and mobile phase B 100% acetonitrile, 0.1% ammonium hydroxide. The TMT labelled

peptide mixture was dissolved in 100μL mobile phase A, centrifuged and injected for fractionation. The gradient elution method at 0.2mL/min included the following steps: 5 minutes isocratic at 5% B, for 35 min gradient to 35% B, gradient to 80% B in 5 min, isocratic for 5 minutes and re-equilibration to 5% B. Signal was recorded at 280nm and fractions were collected every one minute. For cohort 4, peptide fractionation was performed on reversed-phase OASIS HLB cartridges at high pH and up to 9 fractions (10-25% acetonitrile elution steps) were collected for each set. The collected fractions were dried with SpeedVac concentrator and stored at -20°C until LC-MS analysis.

LC-MS analysis

LC-MS analysis was performed on the Dionex Ultimate 3000 UHPLC system coupled with the Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific). Each peptide fraction was reconstituted in 40μL 0.1% formic acid and a volume of 7μL was loaded to the Acclaim PepMap 100, 100μm × 2cm C18, 5μm, 100Å trapping column with the μLPickUp mode at 10μL/min flow rate. The sample was then analysed with a gradient elution on the Acclaim PepMap RSLC (75μm × 50cm, 2μm, 100Å) C18 capillary column retrofitted to an electrospray emitter (New Objective, FS360-20-10-D-20) at 45°C. Mobile phase A was 0.1% formic acid and mobile phase B was 80% acetonitrile, 0.1% formic acid. The gradient method at flow rate 300nL/min was: for 90min gradient to 38% B, for 5min up to 95% B, for 13min isocratic at 95% B, re-equilibration to 5% B in 2min, for 10min isocratic at 10% B. Precursors were selected with 120k mass resolution, AGC 3×10^5 and IT 100ms in the top speed mode within 3sec and were targeted for CID fragmentation with quadrupole isolation width 1.2Th. Collision energy was set at 35% with AGC 1×10^4 and IT 35ms. MS3 quantification spectra were acquired with further HCD fragmentation of the top 10 most abundant CID fragments

isolated with Synchronous Precursor Selection (SPS) excluding neutral losses of maximum m/z 18. Iontrap isolation width was set at 0.7Th for MS1 isolation, collision energy was applied at 55% and the AGC setting was at 6×10^4 with 100ms IT. The HCD MS3 spectra were acquired within 110-400m/z with 60k resolution. Targeted precursors were dynamically excluded for further isolation and activation for 45 seconds with 7ppm mass tolerance. Cohort 4 was analyzed at the MS2 level with a top15 HCD method (CE 40%, 50k resolution) and a maximum precursor intensity threshold of 5×10^7 using the same MS1 parameters as above in a 360min gradient.

Quantification of RNA levels

We used samtools v1.3.1[3] and biobambam v0.0.191[4] to convert cram to fasq files after exclusion of reads that failed QC. We applied FastQC v0.11.5 to check sample quality[5] and excluded 7 samples.

We obtained transcript-level quantification using salmon 0.8.2[6] (with --gcBias and --seqBias flags to account for potential biases) and the GRCh38 cDNA assembly release 87 downloaded from Ensembl [http://ftp.ensembl.org/pub/release-87/fasta/homo_sapiens/cdna/]. We used tximport[7] to convert transcript-level to gene-level scaled transcripts per million (TPM) estimates, with estimates for 39,037 genes based on Ensembl gene IDs.

We excluded 4 samples due to low mapping rate (<80%), 8 samples due to non-European ancestry, 15 samples due to low RIN (<5), 3 samples due to abnormal gene read density plots.

The final gene expression dataset included 87 patients' low-grade and 95 high-grade cartilage samples with 15,249 genes that showed counts per million (CPM) of ≥1 in ≥40 samples (matched low-grade and high-grade samples from 83 patients).

Quantification of protein levels

To carry out protein identification and quantification, we submitted the mass spectra to SequestHT search in Proteome Discoverer 2.1. The precursor mass tolerance was set at 30ppm (Orbitrap Velos data, cohort 1) or 20ppm (Fusion data, cohorts 2-4). For the CID spectra, we set the fragment ion mass tolerance to 0.5Da; for the HCD spectra, to 0.02Da. Spectra were searched for fully tryptic peptides with maximum 2 miss-cleavages and minimum length of 6 amino acids. We specified static modifications as TMT-6-plex at N-terminus, K and Carbamidomethyl at C; dynamic modifications included deamidation of N,Q and oxidation of M. For each peptide, we allowed for a maximum two different dynamic modifications with a maximum of two repetitions. We used the Percolator node to estimate peptide confidence. We set the peptide false discovery rate (FDR) at 1% and based validation on the q-value and decoy database search. We searched all spectra against a UniProt fasta file that contained 20,165 reviewed human entries. The Reporter Ion Quantifier node included a TMT-6-plex (Velos data, cohort 1) or TMT-10-plex (Fusion data, cohorts 2-4) custom Quantification Method with integration window tolerance at 20ppm or 15ppm, respectively. As integration methods, we used the Most Confident Centroid at the MS2 or MS3 level. We only used peptides uniquely belonging to protein groups for quantification.

We excluded samples from 4 patients due to non-European ancestry. The final dataset included low-grade and high-grade cartilage samples each from 99 patients, with 4,801 proteins observed in ≥30% of samples, and 1,677 proteins in all samples, in line with the resolution depth of the isobaric labelling method employed. To account for protein loading, abundance values were normalized by the sum of all protein abundances in a given sample, then log2-transformed and quantile normalized.

Differential RNA expression between high-grade and low-grade cartilage

We tested differential expression of 15,249 genes between high-grade and low-grade cartilage using paired samples from 83 patients. To detect robust gene expression differences, we carried out analyses using different software packages as recommended in a landmark survey of best practices[8], applying limma[9], edgeR[10], and DESeq2[11]. We also tested 5 analysis designs with different options to account for technical variation, including SVAseq[12]. In particular, we tested for differential expression using:

1) A paired analysis of intact and degraded samples (i.e. specifying patient ID as covariate)
2) A paired analysis of intact and degraded samples, with 10 additional covariates accounting for technical variation identified by SVAseq[12]
3) A paired analysis of intact and degraded samples, with 10 RNA sequencing batches as covariates
4) An unpaired analysis of intact and degraded samples
5) An unpaired analysis of intact and degraded samples, with 19 additional covariates accounting for technical variation identified by SVAseq

We tested for differential expression using the following R packages:

I. Limma[9] (with lmFit and eBayes), after applying limma-voom[13] to remove heteroscedasticity
II. DESeq2[11], separately with and without outlier filtering/replacement (minReplicatesForReplace=Inf, cooksCutoff=FALSE options)
III. EdgeR[10], using the likelihood ratio test (glmFit and glmLRT functions), and separately, using the F test (glmQLFit and glmQLFTest functions)

Here and elsewhere, we used Ensembl38p10 to identify genes with uniquely corresponding Ensembl gene ID and gene name (13,737 of 15,249 genes in the RNA data).

In each analysis design and method, we used a 5% False Discovery Rate (FDR) threshold to correct for multiple testing. As the final step, we applied a conservative approach and considered a gene “significantly differentially expressed” between low-grade and high-grade cartilage if it showed significant differential expression across all analysis designs and testing methods (2,557 genes, including 2,418 with uniquely corresponding Ensembl gene ID and gene name).

Differential protein abundance between high-grade and low-grade cartilage

We performed differential analysis for 4,801 proteins that were measured in ≥30% of patients, applying limma[9] to paired samples from 99 patients. Significance was defined at 5% FDR to correct for multiple testing, yielding 2,233 proteins with significant differential abundance (2,019 proteins with uniquely corresponding Ensembl gene ID and gene name).

As batch effects in proteomics data can be pervasive[14,15], paired samples from any patient were always assayed in the same 6-plex (cohort 1) or 10-plex (cohorts 2-4).

Supplementary References

1. Steinberg, J.*, et al.* Integrative epigenomics, transcriptomics and proteomics of patient chondrocytes reveal genes and pathways involved in osteoarthritis. *Sci. Rep.* **7**, 8935 (2017).

2. Steinberg, J.*, et al.* Widespread epigenomic, transcriptomic and proteomic differences between hip osteophytic and articular chondrocytes in osteoarthritis. *Rheumatology (Oxford)* **57**, 1481-1489 (2018).

3. Li, H.*, et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

4. Tischler, G. & Leonard, S. biobambam: tools for read pair collation based algorithms on BAM files. *Source Code Biol. Med.* **9**, 13-13 (2014).

5. Andrews, S. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. (2010).

6. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. salmon provides fast and bias-aware quantification of transcript expression. *Nat. Meth.* **14**, 417-419 (2017).

7. Soneson, C., Love, M. & Robinson, M. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 1; referees: 2 approved]. *F1000Research* **4**, 1521 (2015).

8. Conesa, A.*, et al.* A survey of best practices for RNA-seq data analysis. *Genome Biol.* **17**, 13 (2016).

9. Ritchie, M.E.*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

10. McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288-4297 (2012).

11. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

12. Leek, J.T. svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* **42**, e161 (2014).

13. Law, C.W., Chen, Y., Shi, W. & Smyth, G.K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

14. Gregori, J.*, et al.* Batch effects correction improves the sensitivity of significance tests in spectral counting-based comparative discovery proteomics. *J. Proteom.* **75**, 3938-3951 (2012).

15. Kuligowski, J.*, et al.* Detection of batch effects in liquid chromatography-mass spectrometry metabolomic data using guided principal component analysis. *Talanta* **130**, 442-448 (2014).